

Pred. No. is the number of results predicted by chance to have a

RESULT 1	REFERENCE
AX335784	AUTHOR
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	

AX335784 3866 bp DNA linear PAT 09-JAN-2002
Sequence 6293 from Patent WO0194629.

Homo sapiens
Eumariotyta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D. R. and Weaver, Z.

<p>gene sets</p> <p>Journal Patent: WO 0194629-A 6293 13-DEC-2001; Avalon Pharmaceuticals (US)</p> <p>FEATURES</p> <p>Location/Qualifiers</p> <p>1. 3866</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="unassigned DNA"</p> <p>/db_xref="taxon:9606"</p>									
ORIGIN									
<p>Query Match</p> <p>Best Local Similarity 100.0%; Score 3866; DB 6; Length 3866;</p> <p>Matches 3866; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</p>									
QY	1	GGAAATCCCTTTT	TTTTTTT	TTTGGAGATGGAGTT	CTACTT	GTGTG	CCAGGCTGGAGTG	60	
Db	1	GGAAATCCCTTTT	TTTTTTT	TTTGGAGATGGAGTT	CTACTT	GTGTG	CCAGGCTGGAGTG	60	
QY	61	CAATGGCACAATCT	CAGCTT	CTCAACCTCCGCT	CCCGGTT	CAAGCGAT	CTCCTGC	120	
Db	61	CAATGGCACAATCT	CAGCTT	CTCAACCTCCGCT	CCCGGTT	CAAGCGAT	CTCCTGC	120	
QY	121	CTCAGCTCTCAAGT	AGTGGATT	CAGGCAATGTG	CCACCCT	CGCTG	CTAACTAATTT	180	
Db	121	CTCAGCTCTCAAGT	AGTGGATT	CAGGCAATGTG	CCACCCT	CGCTG	CTAACTAATTT	180	
QY	181	CTTTCTTATTTAGT	AGATGGGTTT	CACATGTTGGT	CAGCTGGT	CTTGA	CTCTG	240	
Db	181	CTTTCTTATTTAGT	AGATGGGTTT	CACATGTTGGT	CAGCTGGT	CTTGA	CTCTG	240	
QY	241	ACCTCAGGTATCACT	TTGGCTT	GGCCTCCCAAGT	CTAGGATT	CAGCCGT	GAAACTG	300	
Db	241	ACCTCAGGTATCACT	TTGGCTT	GGCCTCCCAAGT	CTAGGATT	CAGCCGT	GAAACTG	300	
QY	301	TGCTGTGCTGATCT	TTTTTTT	TTTGGATTTT	TGAAACAGG	GTCTCCCT	TGGTGGCC	360	
Db	301	TGCTGTGCTGATCT	TTTTTTT	TTTGGATTTT	TGAAACAGG	GTCTCCCT	TGGTGGCC	360	
QY	361	AGGCTGGAGTGCAGT	GTGGATCT	TGGCTCACT	TAACTC	CACTCT	CGTTTCAAGT	420	
Db	361	AGGCTGGAGTGCAGT	GTGGATCT	TGGCTCACT	TAACTC	CACTCT	CGTTTCAAGT	420	
QY	421	GATCCTCCCACTTT	AGCTCCT	CTGAGTGTG	TGATTA	CAGGCTG	CAACACCCGG	480	
Db	421	GATCCTCCCACTTT	AGCTCCT	CTGAGTGTG	TGATTA	CAGGCTG	CAACACCCGG	480	
QY	481	CTAATTTTGTATTTT	TTATAGACAGG	GTTCACCATG	TTGGCCAGG	CTGTCT	CTCAA	540	
Db	481	CTAATTTTGTATTTT	TTATAGACAGG	GTTCACCATG	TTGGCCAGG	CTGTCT	CTCAA	540	
QY	541	CTCCTGGACTCAAGG	ATCGCCT	CGCTCCACAT	TCCAAAGT	CCCGAGAT	TACAGTGTG	600	
Db	541	CTCCTGGACTCAAGG	ATCGCCT	CGCTCCACAT	TCCAAAGT	CCCGAGAT	TACAGTGTG	600	
QY	601	AGTCACCATGCTGAC	CTTAATCT	TAACTT	CTAGTCA	TTTCTG	CTCCTTCT	660	
Db	601	AGTCACCATGCTGAC	CTTAATCT	TAACTT	CTAGTCA	TTTCTG	CTCCTTCT	660	
QY	661	GGTCTCCACAACT	CTGATTTAGG	CGGTACAAAT	ATCCTT	AACTT	CATGATTCACAA	720	
Db	661	GGTCTCCACAACT	CTGATTTAGG	CGGTACAAAT	ATCCTT	AACTT	CATGATTCACAA	720	
QY	721	AGGAAGATGAGT	GATTCATG	TTTAAAGGG	GAAGTAG	TAAGCC	CTGCACACTCT	780	
Db	721	AGGAAGATGAGT	GATTCATG	TTTAAAGGG	GAAGTAG	TAAGCC	CTGCACACTCT	780	
QY	781	GGATGATGATCC	TAAATCCAG	TACAGT	TAAAT	TGGGTAT	GGGAAGTAGAATACAAA	840	
Db	781	GGATGATGATCC	TAAATCCAG	TACAGT	TAAAT	TGGGTAT	GGGAAGTAGAATACAAA	840	
QY	841	TTTGGTTTAAATTA	ATCTAATAT	CTAAATA	CTTAA	AAACAT	TTTGGATACATGTTGATGA	900	
Db	841	TTTGGTTTAAATTA	ATCTAATAT	CTAAATA	CTTAA	AAACAT	TTTGGATACATGTTGATGA	900	

QY	901	ATGTAAGACTGTACAG	ACTTCT	CAGAAAA	CAGTTTGG	GTTCAT	CTTTTCA	TTTCCCCAG	960
Db	901	ATGTAAGACTGTACAG	ACTTCT	CAGAAAA	CAGTTTGG	GTTCAT	CTTTTCA	TTTCCCCAG	960
QY	961	TGCAGTTTCTGTAGA	AAATCCG	AGAGATTT	TAAGTGG	CAGAGAA	TGCAAT	TGATT	1020
Db	961	TGCAGTTTCTGTAGA	AAATCCG	AGAGATTT	TAAGTGG	CAGAGAA	TGCAAT	TGATT	1020
QY	1021	CCATAATGAACAA	AGTGTAG	AGACATTT	AAATAAG	TTTAA	AAATGA	AGACCTT	1080
Db	1021	CCATAATGAACAA	AGTGTAG	AGACATTT	AAATAAG	TTTAA	AAATGA	AGACCTT	1080
QY	1081	AACCTAAGCTTGA	ATAAAA	TTTCTG	CTGATCT	TACAGAT	TAAC	CTGGGAA	1140
Db	1081	AACCTAAGCTTGA	ATAAAA	TTTCTG	CTGATCT	TACAGAT	TAAC	CTGGGAA	1140
QY	1141	TTATGATGATGCA	AAACCA	CCAGAG	AGTGG	TTGAGT	TTGCT	CAAACTAG	1200
Db	1141	TTATGATGATGCA	AAACCA	CCAGAG	AGTGG	TTGAGT	TTGCT	CAAACTAG	1200
QY	1201	ACAGTGTTCGCT	TAAGT	GATGCT	CTTTTAA	ATAA	TGAT	TGGT	1260
Db	1201	ACAGTGTTCGCT	TAAGT	GATGCT	CTTTTAA	ATAA	TGAT	TGGT	1260
QY	1261	TTGAAGCGCTT	CCCCC	CAGAT	TAATA	TGGCC	AAAT	TGAG	1320
Db	1261	TTGAAGCGCTT	CCCCC	CAGAT	TAATA	TGGCC	AAAT	TGAG	1320
QY	1321	GATTTGCTGAAT	TAAAG	CTATT	CAAG	AGCCAG	ATGAT	GCAG	1380
Db	1321	GATTTGCTGAAT	TAAAG	CTATT	CAAG	AGCCAG	ATGAT	GCAG	1380
QY	1381	CCAGAGCAAA	CTGCA	AGAA	TTTCT	TTTGT	CTAT	ATAT	1440
Db	1381	CCAGAGCAAA	CTGCA	AGAA	TTTCT	TTTGT	CTAT	ATAT	1440
QY	1441	CACAAAGT	TAATG	CAAAAA	AGTAA	CAAA	CTTCT	TTCA	1500
Db	1441	CACAAAGT	TAATG	CAAAAA	AGTAA	CAAA	CTTCT	TTCA	1500
QY	1501	TACCACTAG	AAATCT	CGGAA	TTT	TAA	ACCTT	CA	1560
Db	1501	TACCACTAG	AAATCT	CGGAA	TTT	TAA	ACCTT	CA	1560
QY	1561	TTTCAGAGG	AGAA	AGAA	AGAA	AGAA	AGAA	AGAA	1620
Db	1561	TTTCAGAGG	AGAA	AGAA	AGAA	AGAA	AGAA	AGAA	1620
QY	1621	TTTCAGAGG	AGAA	AGAA	AGAA	AGAA	AGAA	AGAA	1680
Db	1621	TTTCAGAGG	AGAA	AGAA	AGAA	AGAA	AGAA	AGAA	1680
QY	1681	CTACTAAG	CCAGG	TTT	TAT	TGAG	AGAA	CAAT	1740
Db	1681	CTACTAAG	CCAGG	TTT	TAT	TGAG	AGAA	CAAT	1740
QY	1741	ACCGGAAT	TCAT	TGAG	AGAA	CAAT	TAA	AGT	1800
Db	1741	ACCGGAAT	TCAT	TGAG	AGAA	CAAT	TAA	AGT	1800
QY	1801	CAGTTAAC	CTT	TAAT	TAG	CCAG	AT	TG	1860
Db	1801	CAGTTAAC	CTT	TAAT	TAG	CCAG	AT	TG	1860
QY	1861	GTTTTATG	AAAA	GAG	CAAA	CCCT	CTAG	AT	1920
Db	1861	GTTTTATG	AAAA	GAG	CAAA	CCCT	CTAG	AT	1920
QY	1921	AACCAAGT	GAAAT	GA	TTCT	CTG	TAA	AGT	1980
Db	1921	AACCAAGT	GAAAT	GA	TTCT	CTG	TAA	AGT	1980

MEDLINE PUBMED COMMENT FEATURES source	92348472 1639825 Original 1..3866 Location/Qualifiers /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /cell_line="Yt2C2" /cell_type="T cell" /germline 1..3866 /gene="TTK" 1..979 /note="putative" 1026..3551 /gene="TTK" /codon_start=1 /product="kinase" /protein_id="AAA61239.1" /db_xref="GI:340011" /translation="MNKVRDIKNKFNEDLTDELNKLISADTTDINSVTNVOIMKMAN NPREDLSLLKLEKNSVPLSDALLNKLIGYSQALIEALPDPKYQONESPARIQVFAE LKAIEPDDARDYFQMARANKKFAFVHISFAQFELSQGNVAKSKOLLQKAVRGAAP LEMLEIALNLNLKQKOLLSEBEKKNLASVLTQAESFGSLGHLNRRNSCDRGP TTKARFVIGENMPQDAEIGYRNSLRQNTKQSCFPGRVFNLLNSPDCVKTDDSV VPCFMKQTSRCDRLVPGSPGNSCELRLNLSVQNSHPFKEPLVSDKSSLLI TDSITLKNKTESILLAKLESTKEYQEPVESNOKOWAKRSECIENONPAASSNHQ IPELARKVNTQKHTTPEQPVFSVKQSPPISTKWFDPKSIKCTPSNTLDDVMSCF RTPVKNDRPAPCOLSTPGXPACFOQOQOHLATPLQNLQVLASSANECSIVKRI YSILKQISGSSKSVQVLNEKQIIAIFKYNLEADNQLDSIRNEIAYLNKLQOHS DKLRLDYITQYIYNVECCNIDNLNWLKKSIDPWRKSYMNKMLEAVHTIHQ HGIHSDLPKAFELVDGMLKLIDFGIANQMPDTSVVKDSQVGTVMYMPPEAIKDM SSRENGKSKSIKPSKSDVMSLGLIYMTYKTPQQIINGISKLHALIDPNHIEF PDIPKLODLVCLLRDPKQISIPPELLAHVYVOIOTHPVNMAKGTTEEMKYVLG QLVNLSPNSILKAANTLYEHYSGSHSSSSKTEKKRKK" 1026..2550 misc_feature /gene="TTK" /note="regulatory or binding region; putative" 2551..3351 /gene="TTK" /note="kinase domain; putative" 3353..3866 /gene="TTK" /note="putative" 3843..3848 /gene="TTK" /note="putative"	Db	241	ACCTCAGGTGATCCACATTCGCTTGGCCTCCCAAAGTGTAGATTACAGCGTGAACCTG	300
		Qy	301	TGCCTGGCTGATCTCTTTTGTGTTTGAATAACAGGGTCTCCCTTGGTCGCC	360
		Db	301	TGCCTGGCTGATCTCTTTTGTGTTTGAATAACAGGGTCTCCCTTGGTCGCC	360
		Qy	361	AGGCTGGAGTGCAGTGGTGGATCTTGCTCAGTATACCTCCACCTCCTGGTTCAAGT	420
		Db	361	AGGCTGGAGTGCAGTGGTGGATCTTGCTCAGTATACCTCCACCTCCTGGTTCAAGT	420
		Qy	421	GATCCTCCCATTTAGCCTCTCTAGTAGCTGTGATTACAGCGTGCACACACCCGG	480
		Db	421	GATCCTCCCATTTAGCCTCTCTAGTAGCTGTGATTACAGCGTGCACACACCCGG	480
		Qy	481	CTAATTTTGTATTTTATTAGACAGGGTTTCCACCATTTGGCCAGGCTGTTCTCAA	540
		Db	481	CTAATTTTGTATTTTATTAGACAGGGTTTCCACCATTTGGCCAGGCTGTTCTCAA	540
		Qy	541	CTCCTGGAATCAAGGGATCCGCTCCCTCCACTTCCCAAAGTCCCGAGATTACAGGTGTG	600
		Db	541	CTCCTGGAATCAAGGGATCCGCTCCCTCCACTTCCCAAAGTCCCGAGATTACAGGTGTG	600
		Qy	601	AGTACCATGCTGACCTTATTAATCTTAAGTCAATTTTCTGGTCCATTTCTTCCTTAG	660
		Db	601	AGTACCATGCTGACCTTATTAATCTTAAGTCAATTTTCTGGTCCATTTCTTCCTTAG	660
		Qy	661	GGTCTCTCAACAAATCTGCATTAGGCGGTACAATAATCTTAACTTCAATGATTACAAA	720
		Db	661	GGTCTCTCAACAAATCTGCATTAGGCGGTACAATAATCTTAACTTCAATGATTACAAA	720
		Qy	721	AGGAAGATGAAGTGAATCATGATTGAAAGGGGAAGTAGTAAGCCACCTGCACTCTCT	780
		Db	721	AGGAAGATGAAGTGAATCATGATTGAAAGGGGAAGTAGTAAGCCACCTGCACTCTCT	780
		Qy	781	GGATGATGATCCTAAATCCAGATACAGTAAATAGGGGTATGGGAAGTGAATACAAA	840
		Db	781	GGATGATGATCCTAAATCCAGATACAGTAAATAGGGGTATGGGAAGTGAATACAAA	840
		Qy	841	TTTGGTTTAAATTAATTAATCTAAATATCTAAATATCTAAATATCTAAATATCTAAAT	900
		Db	841	TTTGGTTTAAATTAATTAATCTAAATATCTAAATATCTAAATATCTAAATATCTAAAT	900
		Qy	901	ATGTAAGACTGTACAGACTTCTAGTAAACAGTGGTTCCTATTTTCAATTTCCCTCAG	960
		Db	901	ATGTAAGACTGTACAGACTTCTAGTAAACAGTGGTTCCTATTTTCAATTTCCCTCAG	960
		Qy	961	TGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGATTGACAATGATT	1020
		Db	961	TGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGATTGACAATGATT	1020
		Qy	1021	CCATAATGAACAAAGTGAAGACATTAATAATTAAGTTTAAATAATGAAGCTTACTGATG	1080
		Db	1021	CCATAATGAACAAAGTGAAGACATTAATAATTAAGTTTAAATAATGAAGCTTACTGATG	1080
		Qy	1081	AACTAAAGCTTGAATTAATAATTTCTGCTGATACAGTAACTCGGAACTGTTTAAACAA	1140
		Db	1081	AACTAAAGCTTGAATTAATAATTTCTGCTGATACAGTAACTCGGAACTGTTTAAACAA	1140
		Qy	1141	TTATGATGATGGCAACCAACCCAGAGAGCTGGTGGTGGTTCCTCAAACTAGAGAAA	1200
		Db	1141	TTATGATGATGGCAACCAACCCAGAGAGCTGGTGGTGGTTCCTCAAACTAGAGAAA	1200
		Qy	1201	ACAGTGTTCGCTAAAGTGTGCTCTTTTAAATAATTAAGTTTGGTTCCTCAAGCAAA	1260
		Db	1201	ACAGTGTTCGCTAAAGTGTGCTCTTTTAAATAATTAAGTTTGGTTCCTCAAGCAAA	1260
		Qy	1261	TTGAAGCGCTTCCCGCAGATTAATATGGCCAAATGAGAGTTTGTCTAGAAATCAAGTGA	1320
		Db	1261	TTGAAGCGCTTCCCGCAGATTAATATGGCCAAATGAGAGTTTGTCTAGAAATCAAGTGA	1320
		Qy	1321	GATTTCTGATTAATAAGCTATTCAGAGCCGATGATGACGCTGACTACTTTCAAACTGG	1380
		Db	1321	GATTTCTGATTAATAAGCTATTCAGAGCCGATGATGACGCTGACTACTTTCAAACTGG	1380

QY	1381	CCAGAGCAAACTGCAAGAAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAACTGT	1440
DB	1381	CCAGAGCAAACTGCAAGAAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAACTGT	1440
QY	1441	CACNAGGTATGTCAAAAAAGTAAACAACTTCTTCAAAAAGCTGTAGAACGTGGAGCAG	1500
DB	1441	CACNAGGTATGTCAAAAAAGTAAACAACTTCTTCAAAAAGCTGTAGAACGTGGAGCAG	1500
QY	1501	TACCACATAGAAAATGCTGGAAATTTGCCCTGGGAAATTTAAACCTCCAAAAAAGCAGCTGC	1560
DB	1501	TACCACATAGAAAATGCTGGAAATTTGCCCTGGGAAATTTAAACCTCCAAAAAAGCAGCTGC	1560
QY	1561	TTTCAGAGGAGGAAAAAGAGAAATTTATNCAGATCTACGGTATTAACTGCCCAAGAATCAT	1620
DB	1561	TTTCAGAGGAGGAAAAAGAGAAATTTATNCAGATCTACGGTATTAACTGCCCAAGAATCAT	1620
QY	1621	TTTCCGGTTCACATTGGGCATTTACAGAAATAGGAACAACAGTTGTGATTCAGAGGACAGA	1680
DB	1621	TTTCCGGTTCACATTGGGCATTTACAGAAATAGGAACAACAGTTGTGATTCAGAGGACAGA	1680
QY	1681	CTACTAAAGCCAGGTTTTTATATGGAGAGAACATGCCACACAGATGCGAGAAATAGGTT	1740
DB	1681	CTACTAAAGCCAGGTTTTTATATGGAGAGAACATGCCACACAGATGCGAGAAATAGGTT	1740
QY	1741	ACCGGAATTCATTGAGACAAACTAAACAAACTAAACAGTCATGCCCATTTTGGAAAGAGTCC	1800
DB	1741	ACCGGAATTCATTGAGACAAACTAAACAAACTAAACAGTCATGCCCATTTTGGAAAGAGTCC	1800
QY	1801	CAGTTAACCTTCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT	1860
DB	1801	CAGTTAACCTTCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT	1860
QY	1861	GTTTTATGAAAAGACAAACCTCTAGATCAGAAATGCCAGAGATTTGGTTGTGCTCGGATCTA	1920
DB	1861	GTTTTATGAAAAGACAAACCTCTAGATCAGAAATGCCAGAGATTTGGTTGTGCTCGGATCTA	1920
QY	1921	AACCAAGTGGAAATGATTCCTGTGAAATTAAGAAATTTAAAGTCGTGTTCAAAAATAGTCATT	1980
DB	1921	AACCAAGTGGAAATGATTCCTGTGAAATTAAGAAATTTAAAGTCGTGTTCAAAAATAGTCATT	1980
QY	1981	TCAAGGAACCTCTGGTGTGATGCAAAAGAGTTCTCGAACTTATTATTACTGATTCAAATAA	2040
DB	1981	TCAAGGAACCTCTGGTGTGATGCAAAAGAGTTCTCGAACTTATTATTACTGATTCAAATAA	2040
QY	2041	CCCTGAAGAAATAAAACCGAATCAAGTCCTTAGCTTAAATTTAGAAGAACTTAAGAGATATC	2100
DB	2041	CCCTGAAGAAATAAAACCGAATCAAGTCCTTAGCTTAAATTTAGAAGAACTTAAGAGATATC	2100
QY	2101	AAGAACACAGAGGTTCCAGAGAGTAGTAACACAGAAACAGTGGCAAGCTAAAGAGAAAGTCAGAGT	2160
DB	2101	AAGAACACAGAGGTTCCAGAGAGTAGTAACACAGAAACAGTGGCAAGCTTAAGAGAAAGTCAGAGT	2160
QY	2161	GTATTAAACCGAATCCTGCTGCATCTTTAAATCACTGGCAGATTCGCGAGTTAGCCCGAA	2220
DB	2161	GTATTAAACCGAATCCTGCTGCATCTTTAAATCACTGGCAGATTCGCGAGTTAGCCCGAA	2220
QY	2221	AAGTTAATACAGACAGAAACATACCACTTTTGTAGCAACCTGCTCTTTTTCAGTTTCAAAAC	2280
DB	2221	AAGTTAATACAGACAGAAACATACCACTTTTGTAGCAACCTGCTCTTTTTCAGTTTCAAAAC	2280
QY	2281	AGTCACCAACCAATATCAACATCTAAATGGTTTGTGCCCAAAATCTATTGTGAAGACACAA	2340
DB	2281	AGTCACCAACCAATATCAACATCTAAATGGTTTGTGCCCAAAATCTATTGTGAAGACACAA	2340
QY	2341	GCAGCAATACCTTCGGATGATTACATGAGCTGTTTTTAGAAGCTCCAGTTGTGAAGAAATGACT	2400
DB	2341	GCAGCAATACCTTCGGATGATTACATGAGCTGTTTTTAGAAGCTCCAGTTGTGAAGAAATGACT	2400
QY	2401	TTCCACCTGCTGTTCAGTTGTCAACACCTTATGSCCAACTGCTGTTTCCAGCAGCAAC	2460
DB	2401	TTCCACCTGCTGTTCAGTTGTCAACACCTTATGSCCAACTGCTGTTTCCAGCAGCAAC	2460

Qy	2461	AGCATCAAAATAC	TTGCCAC	TCCACTT	CAAAAATTT	TACAGG	TTTTAG	CAATCTT	CTTTCAGCAA	2520
Db	2461	AGCATCAAAATAC	TTGCCAC	TCCACTT	CAAAAATTT	TACAGG	TTTTAG	CAATCTT	CTTTCAGCAA	2520
Qy	2521	ATGAATGCATTT	CGGTTAA	AGGAAG	AATTTAT	TCCCAT	TATTA	AAAGCAG	ATAGGAAGTGGAG	2580
Db	2521	ATGAATGCATTT	CGGTTAA	AGGAAG	AATTTAT	TCCCAT	TATTA	AAAGCAG	ATAGGAAGTGGAG	2580
Qy	2581	GTTCAAGCAAGG	TATTT	CAGGTGT	TAAATG	AAAAAG	AAAA	CAGATAT	ATATGCTTAT	2640
Db	2581	GTTCAAGCAAGG	TATTT	CAGGTGT	TAAATG	AAAAAG	AAAA	CAGATAT	ATATGCTTAT	2640
Qy	2641	TGAAC	TTTAGA	AGACG	AGATAA	CCAA	ACTCTT	GATAGT	TACCGGAAG	2700
Db	2641	TGAAC	TTTAGA	AGACG	AGATAA	CCAA	ACTCTT	GATAGT	TACCGGAAG	2700
Qy	2701	TGAATAAACT	TACAA	CAACAC	CAGTGA	TAAGAT	CA	TCCGAC	TTTATG	2760
Db	2701	TGAATAAACT	TACAA	CAACAC	CAGTGA	TAAGAT	CA	TCCGAC	TTTATG	2760
Qy	2761	ACCA	GTACATCTA	CATGGT	TATG	AGGTG	TGGAA	ATATTT	TGATCTT	2820
Db	2761	ACCA	GTACATCTA	CATGGT	TATG	AGGTG	TGGAA	ATATTT	TGATCTT	2820
Qy	2821	AGAAAAAT	CCAT	TG	ATCC	ATG	CGAAG	AGTT	TACTG	2880
Db	2821	AGAAAAAT	CCAT	TG	ATCC	ATG	CGAAG	AGTT	TACTG	2880
Qy	2881	TTCA	CAACA	ATCC	AT	CA	ATG	CGA	ATCTT	2940
Db	2881	TTCA	CAACA	ATCC	AT	CA	ATG	CGA	ATCTT	2940
Qy	2941	TAGTTG	ATGGA	ATC	CTA	AA	CTA	TTT	TG	3000
Db	2941	TAGTTG	ATGGA	ATC	CTA	AA	CTA	TTT	TG	3000
Qy	3001	CAACA	AGTGT	TG	TTAA	AG	AT	CT	CAG	3060
Db	3001	CAACA	AGTGT	TG	TTAA	AG	AT	CT	CAG	3060
Qy	3061	TCAA	GATAT	CT	CTT	CC	AG	AGAA	TCT	3120
Db	3061	TCAA	GATAT	CT	CTT	CC	AG	AGAA	TCT	3120
Qy	3121	GTG	ATG	TTT	TG	CT	CT	TAG	ATG	3180
Db	3121	GTG	ATG	TTT	TG	CT	CT	TAG	ATG	3180
Qy	3181	AGC	AGATA	AT	TA	AT	CA	GT	CA	3240
Db	3181	AGC	AGATA	AT	TA	AT	CA	GT	CA	3240
Qy	3241	AAT	TTCC	GA	TAT	TC	AG	AGAA	CA	3300
Db	3241	AAT	TTCC	GA	TAT	TC	AG	AGAA	CA	3300
Qy	3301	ACCC	AAAA	CAG	AG	GAT	AT	CC	AT	3360
Db	3301	ACCC	AAAA	CAG	AG	GAT	AT	CC	AT	3360
Qy	3361	CTCA	TC	AG	TTAA	CC	AA	AT	GC	3420
Db	3361	CTCA	TC	AG	TTAA	CC	AA	AT	GC	3420
Qy	3421	AAC	TTG	TTG	GT	CT	CT	AA	CT	3480
Db	3421	AAC	TTG	TTG	GT	CT	CT	AA	CT	3480
Qy	3481	ACT	AT	AG	TTG	GT	CT	CT	AA	3540
Db	3481	ACT	AT	AG	TTG	GT	CT	CT	AA	3540
Qy	3541	GA	AAAA	AT	GA	TT	TC	CA	AT	3600

Db 3541 GAAAAAATGATTTGCGATTATTCGTAATGTCAGATAGGAGGTATAAAATATATGGACT 3600
 QY 3601 GTTATCTCTTGAATCCCTGCGAAATCTACATTTGAAGACAACATCACTCTGAAGTGT 3660
 Db 3601 GTTATCTCTTGAATCCCTGCGAAATCTACATTTGAAGACAACATCACTCTGAAGTGT 3660
 QY 3661 ATCAGCAAAAAAATTCAGTGAGATTAATCTTTAAAGAAAACTGTAAAAATAGCAACAC 3720
 Db 3661 ATCAGCAAAAAAATTCAGTGAGATTAATCTTTAAAGAAAACTGTAAAAATAGCAACAC 3720
 QY 3721 TTATGGCACTGTATATTCAGTCTGTTTCTCTCTTTTATGCTCTGTTGTAATCTAC 3780
 Db 3721 TTATGGCACTGTATATTCAGTCTGTTTCTCTCTTTTATGCTCTGTTGTAATCTAC 3780
 QY 3781 TTGACATCAATTTACTCTTGGATAGTGGTGGATAGCAAGTATATTTAAAAAACTTTG 3840
 Db 3781 TTGACATCAATTTACTCTTGGATAGTGGTGGATAGCAAGTATATTTAAAAAACTTTG 3840
 QY 3841 TAAATAAGTTTGTGCTAAATGA 3866
 Db 3841 TAAATAAGTTTGTGCTAAATGA 3866

RESULT 3
 LOCUS BC032858 2980 bp mRNA linear PRI 12-NOV-2003
 DEFINITION Homo sapiens TTK protein kinase, mRNA (cdna clone MGC:33286
 IMAGE:527001.2), complete cds.
 ACCESSION BC032858
 VERSION BC032858.2 GI:34190560
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2980)
 AUTHORS Klausner, R.D., Collins, F.S., Wagner, L., Shennen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Wax, S.L., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loguercio, N.A., Peters, G.J.,
 Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.O. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 MEDLINE 22388257
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 2980)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (07-JUN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK NTH-MGC Project URL: <http://mhc.nci.nih.gov>
 On Aug 25, 2003 this sequence version replaced gi:23271248.
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadan@systemsbiology.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

FEATURES

Location/Qualifiers
 1..2980
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:33286 IMAGE:527001.2"
 /tissue_type="Testis"
 /clone_lib="NIH MGC_97"
 /lab_host="DH10B"
 /note="Vector: pBluescript"
 1..2980
 /gene="TTK"
 /note="synonyms: MPS1L1, PYT, ESK"
 /db_xref="LocusID:7272"
 /db_xref="MIM:604092"
 77..2650
 /codon_start=1
 /product="TTK protein kinase"
 /protein_id="AAH32858.1"
 /db_xref="GI:23271249"
 /db_xref="LocusID:7272"
 /translation="MESEDSGRELTIDSIMKVRDIKNKFNEDLTDELSLNKISAD
 TTDNSGTNQIMMMNPEDWLSLLKLEKNSVPLSDALLNKILIGYSQAIAPDPK
 YGQNSFARQVPAELKAQEPDARDYFQMARCKKFAFVHISPAQFAELSGNVK
 KSKOLLOKVERGAPLEMLEALENLNLOKOLLSEKKNSASTVLTAFSGVSG
 LGLHONRNSCDSSGQTKARFLYGENMPQDAEIGYRNSLRQTNKTKQSCPRGVPV
 NLNPSDCVDKDDSVVPCFMKQTSRSECDLVVPGSGESGNSCELRLKSKVQNSH
 FKPELVDEKSSLLIITDSITLKNKTESLLAKLEETKEQEPVESPNQKQVSKRK
 SECINQNPAAASNNHQIPELARKVTEQKHTTPEQPVFSKQSPISSTSKWDFPKI
 CKTPSNTLDYMSCFRTVVRKNDPPACQLSTPQVFPVSPQACQCCQIILATLQNLQV
 LASSANECSVKGRIYSILKQIGSGSSKVFOLNEKKOYIAKYVNLDEADNQLD
 SYRNEIAYLANKLOQHSKPIRLVDYEITDOYIYMWECGNIDINSWLKKKSIDPWER
 KSYHKNMLEAVHIHQHGIHVHSLKPAFLIVGMLKLIDFGIANQMPDTTSVVDK
 QVGIVNTPPPEAKDMSSRENGSKSKISPKSDVSLGCLILYMTYKTPFQOINQ
 ISKLHAIDPNHEIFDPEKLDQVLKCKLRDPKORISIPPELLAHVPVQIOTHPV
 NQWAKGTTEEMKTVLGQVLGNSPNSILKAAKTLYEHYSGSGSHSSSKTFEKGK
 K"
 misc_feature 1649..2449
 /note="S_TKC; Region: Serine/Threonine protein kinases,
 catalytic domain"
 /db_xref="CDD:smart00220"

ORIGIN

Query Match 75.4%; Score 2914; DB 9; Length 2980;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 2931; Conservative 0; Mismatches 10; Indels 1; Gaps 1;
 QY 925 GAAACAGTTTGGGTTCCATCTTTTCATTTCCCGAGTCAGTTCCTGTAGAATGGAAT 984
 Db 24 GAAAGAGTTTGGGTTCCATCTTTTCATTTCCCGAGTCAGTTCCTGTAGAATGGAAT 83
 QY 985 CCGAGGATTTAAGTGGCAGAGATTGACATTCGATCCATAATCAACAAAGTCAGAGACA 1044
 Db 84 CCGAGGATTTAAGTGGCAGAGATTGACATTCGATCCATAATCAACAAAGTCAGAGACA 143
 QY 1045 TTAATAAAGTTTAAAAAATGAAGACCTTACTGATGAACCTTAAGCTTAATAAATTTCTG 1104
 Db 144 TTAATAAAGTTTAAAAAATGAAGACCTTACTGATGAACCTTAAGCTTAATAAATTTCTG 203

QY 1105 CTGATACTACAGATAACTCGGGAACCTGTTAAACCAAAATTATGATGATGGCAAAACCAACCCAG 1164
DB 204 CTGATACTACAGATAACTCGGGAACCTGTTAAACCAAAATTATGATGATGGCAAAACCAACCCAG 263
QY 1165 AGGACTGGTGGAGTTGCTGCTCAAACTAGAGAAAACAGTGTTCGGTAAAGTGAATGCTC 1224
DB 264 AGGACTGGTGGAGTTGCTGCTCAAACTAGAGAAAACAGTGTTCGGTAAAGTGAATGCTC 323
QY 1225 TTTTAAATAAATTGATTGGTGGTTTACAGTCAAGCAATTGAAGCGCTTCCGCCACAGATAAAT 1284
DB 324 TTTTAAATAAATTGATTGGTGGTTTACAGTCAAGCAATTGAAGCGCTTCCGCCACAGATAAAT 383
QY 1285 ATGGCCAAAATAGAGAGTTTGGCTAGAAATCAAGTGAGATTTGCTGTAATTAAGAGCTATTTC 1344
DB 384 ATGGCCAAAATAGAGAGTTTGGCTAGAAATCAAGTGAGATTTGCTGTAATTAAGAGCTATTTC 443
QY 1345 AAGAGCCAGATGATGCACGTGACTTCTCAAAATGGCCAGAGCAAACTGCAAGAAATTTG 1404
DB 444 AAGAGCCAGATGATGCACGTGACTTCTCAAAATGGCCAGAGCAAACTGCAAGAAATTTG 503
QY 1405 CTTTGTTCATATCTTTTGCACAAATTTGAACCTGTCACAGGTAAATGTCAAAAAAGTA 1464
DB 504 CTTTGTTCATATCTTTTGCACAAATTTGAACCTGTCACAGGTAAATGTCAAAAAAGTA 563
QY 1465 AACCACTTCTTCAAAAAGCTGTAGAACGTGGAGCAGTACCCTAGAAATGCTGGAAATTC 1524
DB 564 AACCACTTCTTCAAAAAGCTGTAGAACGTGGAGCAGTACCCTAGAAATGCTGGAAATTC 623
QY 1525 CCCTGCGGAATTTAAACCTCCAAAAAGCAGCTGCTTTCAGAGGAGGAAAAAGAAAT 1584
DB 624 CCCTGCGGAATTTAAACCTCCAAAAAGCAGCTGCTTTCAGAGGAGGAAAAAGAAAT 683
QY 1585 TATCAGCATCTACGGTATTAACCTCCCAAGAAATCAATTTCCGGTTCACITGGGCATTTAC 1644
DB 684 TATCAGCATCTACGGTATTAACCTCCCAAGAAATCAATTTCCGGTTCACITGGGCATTTAC 743
QY 1645 AGAATAGCAACACAGTTGTGATTCACAGAGCAGACTACTTAAAGCCAGAGTTTATATG 1704
DB 744 AGAATAGCAACACAGTTGTGATTCACAGAGCAGACTACTTAAAGCCAGAGTTTATATG 803
QY 1705 GAGAGACATGCCACCAAGATCGAATAATAGTTACCGGAATTCATTGAGCAAACTA 1764
DB 804 GAGAGACATGCCACCAAGATCGAATAATAGTTACCGGAATTCATTGAGCAAACTA 863
QY 1765 ACRAAATTAACAGTGCATGCCATTTGGAAGAGTCCAGATTAACCTTCTTAAATAGCCAG 1824
DB 864 ACRAAATTAACAGTGCATGCCATTTGGAAGAGTCCAGATTAACCTTCTTAAATAGCCAG 923
QY 1825 ATTGTGATGTAAGACAGATGATTTCAGTTGTACCTTGTGTTTATGAAAAGCAAAACCTCTA 1884
DB 924 ATTGTGATGTAAGACAGATGATTTCAGTTGTACCTTGTGTTTATGAAAAGCAAAACCTCTA 983
QY 1885 GATCAGAAATGCCAGATTTGGTTGGCTGGATCTAAACCAAGTGGAAATGATTCCTGTG 1944
DB 984 GATCAGAAATGCCAGATTTGGTTGGCTGGATCTAAACCAAGTGGAAATGATTCCTGTG 1043
QY 1945 AATTAAGAAATTTAAAGTCTCTTCAAAATAGTCAATTTCAAGAACCTCTGTTGTCAGATG 2004
DB 1044 AATTAAGAAATTTAAAGTCTCTTCAAAATAGTCAATTTCAAGAACCTCTGTTGTCAGATG 1103
QY 2005 AAAAGAGTCTGAACCTTATTAATTAATGATTCATTAACCTGGAAGATAAAACCGGAATCAA 2064
DB 1104 AAAAGAGTCTGAACCTTATTAATTAATGATTCATTAACCTGGAAGATAAAACCGGAATCAA 1163
QY 2065 GTCTTCTAGCTAAATTAAGAAACCTTAAAGAGTATCAAGACACAGAGGTTCCAGAGAGTA 2124
DB 1164 GTCTTCTAGCTAAATTAAGAAACCTTAAAGAGTATCAAGAAACAGAGGTTCCAGAGAGTA 1223
QY 2125 ACCAGAAACAGTGGCAAGCTTAAGAGAAAGTCAAGAGTGTATTAAACAGAACTCTGCTGCAT 2184
DB 1224 ACCAGAAACAGTGGCAATCTTAAGAGAAAGTCAAGAGTGTATTAAACAGAACTCTGCTGCAT 1283
QY 2185 CTTCAATCACTGGCAGATTCGGAGTTAGCCCAAAAAGTTAATACAGAGCAGAAACATA 2244

DB 1284 CTTCAATCACTGGCAGATTCGGAGTTAGCCCAAAAAGTTAATACAGAGCAGAAACATA 1343
QY 2245 CCACITTTTGAGCAACCTGCTTTTTCAGTTTCAAAAACAGTCCACCAATATCAACATCTA 2304
DB 1344 CCACITTTTGAGCAACCTGCTTTTTCAGTTTCAAAAACAGTCCACCAATATCAACATCTA 1403
QY 2305 AATGTTTGAACCCAAAATCTATTTGTAAGACACCAAGCAGCAATACCTTGGATGATTACA 2364
DB 1404 AATGTTTGAACCCAAAATCTATTTGTAAGACACCAAGCAGCAATACCTTGGATGATTACA 1463
QY 2365 TGAGCTGTTTGTAGAACTCCAGTGTAAAGAAATGACTTTCACCTGCTTGTGATGTTGCTCAA 2424
DB 1464 TGAGCTGTTTGTAGAACTCCAGTGTAAAGAAATGACTTTCACCTGCTTGTGATGTTGCTCAA 1523
QY 2425 CACCTTATGGCCAAACCTGCTGTTTTCAGAGCAGCAACAGCATCAATACTTGCACCTCAC 2484
DB 1524 CACCTTATGGCCAAACCTGCTGTTTTCAGAGCAGCAACAGCATCAATACTTGCACCTCAC 1583
QY 2485 TTCAAAAATTTACAGGTTTTAGCATCTTCTTCAAGCAATGAATGCAATTCGGTTAAAGGAA 2544
DB 1584 TTCAAAAATTTACAGGTTTTAGCATCTTCTTCAAGCAATGAATGCAATTCGGTTAAAGGAA 1643
QY 2545 GAATTTATTTCCATATTTAAAGCAGATAGGAAGTGGAGGTTCAAGCAAGTATTTCAAGTGT 2604
DB 1644 GAATTTATTTCCATATTTAAAGCAGATAGGAAGTGGAGGTTCAAGCAAGTATTTCAAGTGT 1703
QY 2605 TAAATGAAAAGAAAACAGATATATGCTATAAATATGTGAACCTTGAAGAAAGCAGATAACC 2664
DB 1704 TAAATGAAAAGAAAACAGATATATGCTATAAATATGTGAACCTTGAAGAAAGCAGATAACC 1763
QY 2665 AAACCTCTTGATGTTTACCGGACGAAATAGCTTATTTGAATAAACTCAACAACACAGTG 2724
DB 1764 AAACCTCTTGATGTTTACCGGACGAAATAGCTTATTTGAATAAACTCAACAACACAGTG 1823
QY 2725 ATAAGATCATCCGACTTTATGATTTATGAATATGAATCATCGGACAGTACATCTACATGTTAATGG 2784
DB 1824 ATAAGATCATCCGACTTTATGATTTATGAATATGAATCATCGGACAGTACATCTACATGTTAATGG 1883
QY 2785 AGTGTGGAATATTCATCTTANTAGTTGGCTTAAAGAAATAATCCATTTGATTCATCGG 2844
DB 1884 AGTGTGGAATATTCATCTTANTAGTTGGCTTAAAGAAATAATCCATTTGATTCATCGG 1943
QY 2845 AACCGAAGAGTTTACTTGGAAAAATATGTTAGAGGAGTTTCAACAATCCATCAACATGGCA 2904
DB 1944 AACCGAAGAGTTTACTTGGAAAAATATGTTAGAGGAGTTTCAACAATCCATCAACATGGCA 2003
QY 2905 TTGTTCCAGAGTCAATCTTAAACCAAGTAACTTTCTGATGTTGATGGAATGCTTAAAGCTAA 2964
DB 2004 TTGTTCCAGAGTCAATCTTAAACCAAGTAACTTTCTGATGTTGATGGAATGCTTAAAGCTAA 2063
QY 2965 TTGATTTTGGGATTCGAACCAAAATGCAACAGATACAAAGTGTGTTTAAAGATTCTC 3024
DB 2064 TTGATTTTGGGATTCGAACCAAAATGCAACAGATACAAAGTGTGTTTAAAGATTCTC 2123
QY 3025 AGGTTGGCAGAGTTAATTTATATGCCCCAGAGCAATCAAGATATGCTTCTCCACAGAG 3084
DB 2124 AGGTTGGCAGAGTTAATTTATATATGCCCCAGAGCAATCAAGATATGCTTCTCCACAGAG 2183
QY 3085 AGAATGGGAAATCTAAGTCAAGATAGCCCCCAAAAGTGTGTTTGGTCTTGGATGTA 3144
DB 2184 AGAATGGGAAATCTAAGTCAAGATAGCCCCCAAAAGTGTGTTTGGTCTTGGATGTA 2243
QY 3145 TTTTGTACTATATGACTTACGGGAAAAACCACTTTTCAAGCAATAATTAATCAGATTTCTA 3204
DB 2244 TTTTGTACTATATGACTTACGGGAAAAACCACTTTTCAAGCAATAATTAATCAGATTTCTA 2303
QY 3205 AATTACATGCCAATTAATGATCTTAATGAAATTTGAAATTTCCCGATATTCAGAGAAAG 3264
DB 2304 AATTACATGCCAATTAATGATCTTAATGAAATTTGAAATTTCCCGATATTCAGAGAAAG 2363
QY 3265 ATCTTCAAGATGTTTAAAGTGTGTTTAAAGGAGCCCAAAACAGAGGATATCATTC 3324

Db 2364 ATCTTCAAGATGCTTAAAGTGTGTTTAAAGAGGACCCAAACACAGAGGATATCCATTC 2423
Qy 3325 CTGAGCTCTCTGGCTCATCCATATGTTTCAAAATTCAAATCATCCAGTTAAACAAATGGCCA 3384
Db 2424 CTGAGCTCTCTGGCTCATCCATATGTTTCAAAATTCAAATCATCCAGTTAAACAAATGGCCA 2483
Qy 3385 AGGGAACACCTGAGAAATGAATATGTTCTGGCCCAACTTGTGCTGAATCTTCCTTA 3444
Db 2484 AGGGAACACCTGAGAAATGAATATGTTCTGGCCCAACTTGTGCTGAATCTTCCTTA 2543
Qy 3445 ACTCCATTTTGAAGAGCTGCTAAACATTTATATGAACACTATAGTGGTGGTGAAGTCATA 3504
Db 2544 ACTCCATTTTGAAGAGCTGCTAAACATTTATATGAACACTATAGTGGTGGTGAAGTCATA 2603
Qy 3505 ATTCTTCATCTCCAGAGCTTTTGAAGAAAGAGGGAAGAAATGATTTGCAAGTTATTC 3564
Db 2604 ATTCTTCATCTCCAGAGCTTTTGAAGAAAGAGGGAAGAAATGATTTGCAAGTTATTC 2663
Qy 3565 GTAATGTCAGATAGGAGTATATAATATATTTGGACTGTTATCTCTTGAATCCCTGTGA 3624
Db 2664 GTAATGTCAGATAGGAGTATATAATATATTTGGACTGTTATCTCTTGAATCCCTGTGA 2723
Qy 3625 AATCTACATTTGAAGCAACATCACTCTGAAGTGTATATCAGCAAAAGAAATTCAGTGAGA 3684
Db 2724 AATCTACATTTGAAGCAACATCACTCTGAAGTGTATATCAGCAAAAGAAATTCAGT-AGA 2782
Qy 3685 TTATCTTTAAAGAAACTGTAAATATAGCAACCACTTATGGCACTGTATATATTTAGA 3744
Db 2783 TTATCTTTAAAGAAACTGTAAATATAGCAACCACTTATGGCACTGTATATATTTAGA 2842
Qy 3745 CTTCCTTTCTCTGTTTATGCTCTTGTGTAATCTTACTTGACATCAITTTACTCTTGGAAAT 3804
Db 2843 CTTCCTTTCTCTGTTTATGCTCTTGTGTAATCTTACTTGACATCAITTTACTCTTGGAAAT 2902
Qy 3805 AGTGGTGGATAGCAAGTATATCTTAAAGAACTTTGTAAATAAGTTTGTGGCTAAAT 3864
Db 2903 AGTGGTGGATAGCAAGTATATCTTAAAGAACTTTGTAAATAAGTTTGTGGCTAAAT 2962
Qy 3865 GA 3866
Db 2963 GA 2964

RESULT 4
AX775733
LOCUS 2975 bp mRNA linear PAT 14-JUL-2003
DEFINITION Sequence 3 from Patent WO03048202.
ACCESSION AX775733
VERSION AX775733.1 GI:32693451
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Matsuda, A. and Muramatsu, S.
NF-kB activating gene
Patent: WO 03048202-A 3 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
Location/Qualifiers
1..2975
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
84..2657
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11566.1"
/db_xref="GI:32693452"
/translation="MESEDSGRBELTIDSIMNKRDIKKNFKNEDLTDELSLNKISAD
TTDNGSFVQIMMANNPEDWLSLLKLEKNSVPLSDALLNKLIGRYSOAIFALPDK
YQNGSFARIQVFAELKAIQEPDARDYFQMARANCKFAFVHISFAQFELSQNVK
KSKQLLQKAVERGAVELEMLEIALRNLNLQKQLLSEEBKNLSASTVLTAQESPSGS

CDS
1..2975
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
84..2657
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11566.1"
/db_xref="GI:32693452"
/translation="MESEDSGRBELTIDSIMNKRDIKKNFKNEDLTDELSLNKISAD
TTDNGSFVQIMMANNPEDWLSLLKLEKNSVPLSDALLNKLIGRYSOAIFALPDK
YQNGSFARIQVFAELKAIQEPDARDYFQMARANCKFAFVHISFAQFELSQNVK
KSKQLLQKAVERGAVELEMLEIALRNLNLQKQLLSEEBKNLSASTVLTAQESPSGS

ORIGIN

Query Match 75.3%; Score 2909.2; DB 6; Length 2975;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2928; Conservative 0; Mismatches 13; Indels 1; Gaps 1;
Qy 925 GAAACACGTTTGGGTTCCATCTTTTCATTTCCCCAGTCGAGTTTCTCTGTAGAAATGGAAT 984
Db 31 GAAAGGAGTTTGGGTTCCATCTTTTCATTTCCCCAGTCGAGTTTCTCTGTAGAAATGGAAT 90
Qy 985 CCGAGGATTTAAGTGGCAGAGAAATTGCAATTTGATTTCCATAATGAACAAAGTGAGAGACA 1044
Db 91 CCGAGGATTTAAGTGGCAGAGAAATTGCAATTTGATTTCCATAATGAACAAAGTGAGAGACA 150
Qy 1045 TTAATAATAGTTTAAATAATGAACCTTACTGATGAACTTAAGCTTGAATATAAATTTCTG 1104
Db 151 TTAATAATAGTTTAAATAATGAACCTTACTGATGAACTTAAGCTTGAATATAAATTTCTG 210
Qy 1105 CTGATACTACAGATAAATCGGGAACCTGTTAAACCAAAATTTATGATGATGCGCAAAACACCCAG 1164
Db 211 CTGATACTACAGATAAATCGGGAACCTGTTAAACCAAAATTTATGATGATGCGCAAAACACCCAG 270
Qy 1165 AGGACTGGTTGAGTTTGTCTCAAACTAGAGAAAAACAGTGTTCCTCAAGTATGATGCTC 1224
Db 271 AGGACTGGTTGAGTTTGTCTCAAACTAGAGAAAAACAGTGTTCCTCAAGTATGATGCTC 330
Qy 1225 TTTTAAATAAATTTGATTTGCTTACAGTCAAGCAATTTGAAGCGCTTCCCCAGATAAAT 1284
Db 331 TTTTAAATAAATTTGATTTGCTTACAGTCAAGCAATTTGAAGCGCTTCCCCAGATAAAT 390
Qy 1285 ATGCCCCAAATGAGAGTTTGTCTAGAATTTCAAGTGAAGTTTGTCTGAATTTAAAGCTATTC 1344
Db 391 ATGCCCCAAATGAGAGTTTGTCTAGAATTTCAAGTGAAGTTTGTCTGAATTTAAAGCTATTC 450
Qy 1345 AAGAGCCAGATGATGACGTCAGTCTACTTTCAAAATGCGCAGAGCAAACTGCAAGAAATTTG 1404
Db 451 AAGAGCCAGATGATGACGTCAGTCTACTTTCAAAATGCGCAGAGCAAACTGCAAGAAATTTG 510
Qy 1405 CTTTGTGTTATATATCTTTTGCAAAATTTGAACCTGTCAGAGGAGGAGGAGGAGGAGGAGT 1464
Db 511 CTTTGTGTTATATATCTTTTGCAAAATTTGAACCTGTCAGAGGAGGAGGAGGAGGAGGAGT 570
Qy 1465 AACAACTTCTTCAAAAGCTGTAGAACCTGGAGCAGTACCACCTAGAAATGCTGGAATTTG 1524
Db 571 AACAACTTCTTCAAAAGCTGTAGAACCTGGAGCAGTACCACCTAGAAATGCTGGAATTTG 630
Qy 1525 CCCTGCGGAATTTAAACCTCCAAAAAGCAGCTGCTTTTCAGAGGAGGAGGAGGAGGAGGAGT 1584
Db 631 CCCTGCGGAATTTAAACCTCCAAAAAGCAGCTGCTTTTCAGAGGAGGAGGAGGAGGAGT 690
Qy 1585 TATCAGCATCTACGGTATTAATCTGCCCAAGAAATCAATTTCCGGTTCTACTTGGGCAATTTAC 1644
Db 691 TATCAGCATCTACGGTATTAATCTGCCCAAGAAATCAATTTCCGGTTCTACTTGGGCAATTTAC 750
Qy 1645 AGAATAGAACCAACAGTTTGTGATTTCCAGAGGACAGACTACTTAAAGCAGGTTTATATG 1704
Db 751 AGAATAGAACCAACAGTTTGTGATTTCCAGAGGACAGACTACTTAAAGCAGGTTTATATG 810
Qy 1705 GAGAGAACATGCCACCAACAGATGCAAGAAATAGTTTACCGGAATTCATTGAGACAAACTA 1764
Db 811 GAGAGAACATGCCACCAACAGATGCAAGAAATAGTTTACCGGAATTCATTGAGACAAACTA 870


```

RESULT 5
BC000633
LOCUS
DEFINITION
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
2984 bp mRNA linear PRI 12-NOV-2003
IMAGE:3343925), complete cds.
ACCESSION
BC000633
VERSION
BC000633.2 GI:33990673
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2984)
Srausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shemmen, C.M., Schuler, G.D.,
Altschul, S.F., Zebberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaefer, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prance, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Aramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.E., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Vallian, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22389257
12477932
2 (bases 1 to 2984)
Srausberg, R.
Direct Submission
Submitted (15-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
On Aug 20, 2003 this sequence version replaced gi:12653696.
Contact: MGC help desk
Email: cgabs-rc@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (JLNLI)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NIHC),
Gaithersburg, Maryland
Web site: http://www.nisn.nih.gov/
Contact: nisc.ncgenheri.nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L. H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/JLNLI at: http://image.llnl.gov
Series: IRAL Plate: 5 Row: d Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched MENA gi: 23308721.
Location/Qualifiers
1..2984
/organism="Homo sapiens"
/mol_type="mRNA"

```

db_xref="taxon:9606"
 /clone="MGC:865 IMAGE:3343925"
 /tissue_type="Eye, retinoblastoma"
 /clone_lib="NIH MGC 16"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 1..2984
 /gene="TTTK"
 /note="synonyms: MPS1L1, PYT, ESK"
 /db_xref="LocusID:7272"
 /db_xref="MIM:604092"
 75..2648
 /codon_start=1
 /product="TTK protein kinase"
 /protein_id="AAH00633.1"
 /db_xref="GI:12653697"
 /db_xref="LocusID:7272"
 /translation="MESEDLSGRELITDSIMNKVRDIKNKFNEDLTDELNLNKISAD
 TTDNSGTNQIMMMANNPDWLSLLKLEKNSVPLSDALLNKILIGYSQAIKALPDK
 YGQNESFARIQVFAELKALQEPDARDYQMARANCKKFAFVHISPAQFELSGQNVK
 KSKOLLQKAVERGAVPLEMLEIALENLLOKOLLSEEEKNLASATVLTAFQSFSGS
 LCHLQNRNSCDNRGQTTKARFLYGENMPQDAEIGYRNSLRQINKTKQSCPGRPVY
 NLLNSPDCDVKTDSDVVPCFMKQTSRSECDRLVFGSKPSGNSCELRLNKSQVNSH
 FKPELVDSKSELIITDSITLKNKTESLLAKLEBETKEYQEPVESNQKQWKRK
 SEICINQPAASNNHQIPELARKVTEQKHTTEQPVFVSQSPISITSKWFDPLSI
 CKTPSNTLDDYNSCRTPVKNDFPACOLSTPYQOPACQOQOQIILATPQNLQV
 LASSANECISVKGRIVSIIKQISGSSSKVFOVNEKKOIVAIKYNLEAENDQLD
 SYNEATAYLAKLQOHSDKIIRLYDYRITQVYVMMECNDILNSMLKKKSIDPWER
 KSYNWMLEAVHTIHQGVHSDLKPAFLIVGMLKLDIFGLANQMPDPTTSVQVDS
 QGVTVNMPLEARDSSSRENGSKSPKSDVNSLGLGILYIMTYGTPPOVILNQ
 ISKLHAIDPNHEIFDPDKOLQVLCCKLRDKPKRISPELLAHYPVQIOTHPV
 NOKAKGTTEEMKTVLGQLVGLNSPILKAAKTLVHYSGESHSSSSKTFKKRGK
 K"
 1647..2447
 /note="S_TKc; Region: Serine/Threonine protein kinases,
 catalytic domain"
 /db_xref="CDD:smart00220"

gene

CDS

misc_feature

ORIGIN

Query Match 75.3%; Score 2909.2; DB 9; Length 2984;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 2928; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 925 GAAACAGTTGGGTTCCATCTTTTCATTTCCCGAGTGCAGTTTCTGTAGAAATGGAAT 984
 DB 22 GAAAGGAGTTGGGTTCCATCTTTTCATTTCCCGAGTGCAGTTTCTGTAGAAATGGAAT 81
 QY 985 CCGAGGATTTAAGTGGCAGAGATTGACAAATTCATTCATTAATGAACAAGTCAGAGACA 1044
 DB 82 CCGAGGATTTAAGTGGCAGAGATTGACAAATTCATTCATTAATGAACAAGTCAGAGACA 141
 QY 1045 TTAATAAAGTTTAAATAATGAAGCTTACTGATGAACCTTGAATTAATAATTTCTG 1104
 DB 142 TTAATAAAGTTTAAATAATGAAGCTTACTGATGAACCTTGAATTAATAATTTCTG 201
 QY 1105 CTGATCTACAGATACTCGGAACTGTAAACAAATTTATGATGATGGCAACCAACCCAG 1164
 DB 202 CTGATCTACAGATACTCGGAACTGTAAACAAATTTATGATGATGGCAACCAACCCAG 261
 QY 1165 AGGACTGTTGAGTTTGTCTCTCAAACTAGAGAAAACAGTGTTCGGCTAAGTGTGCTC 1224
 DB 262 AGGACTGTTGAGTTTGTCTCTCAAACTAGAGAAAACAGTGTTCGGCTAAGTGTGCTC 321
 QY 1225 TTTTAAATAAATGATGGTTCGTTACAGTCAAGCAATTTGAAGCGCTTCCCGCAGATAAT 1284
 DB 322 TTTTAAATAAATGATGGTTCGTTACAGTCAAGCAATTTGAAGCGCTTCCCGCAGATAAT 381
 QY 1285 ATGCCCAAAATGAGAGTTTGTCTAGAAATTCAGTGAAGATTTGCTGAATTAAGAGCTATT 1344
 DB 382 ATGCCCAAAATGAGAGTTTGTCTAGAAATTCAGTGAAGATTTGCTGAATTAAGAGCTATT 441
 QY 1345 AAGAGCAGATGATGACGCTGACTCTTTCAATGCGCAGAGCAAACTGCAAGAAATTTG 1404

Db 442 AAGAGCCAGATGATGCAGTGACTACTTTCAAATGCGCAGAGCAAACTGCAAGAAATTG 501
Qy 1405 CTTTGTGTCATATATCTTTTGCACAAATTGAACTGTCTCAGAGGTAAATGTCAAAAAAGTA 1464
Db 502 CTTTGTGTCATATATCTTTTGCACAAATTGAACTGTCTCAGAGGTAAATGTCAAAAAAGTA 561
Qy 1465 AACAACTTCTTCAAAAAGCTGTAGAACGTGGAGCAGTACCCTAGAAAATGCTGGAATTTG 1524
Db 562 AACAACTTCTTCAAAAAGCTGTAGAACGTGGAGCAGTACCCTAGAAAATGCTGGAATTTG 621
Qy 1525 CCCTGCGGAATTTAAACCTCCAAAAGAGCAGCTCTTTCAGAGGAGGAAGAGAAATTT 1584
Db 622 CCCTGCGGAATTTAAACCTCCAAAAGAGCAGCTCTTTCAGAGGAGGAAGAGAAATTT 681
Qy 1585 TATCAGCATCTACCGGTATTAACCTGCCCAAGAAATCATTTCCGGTTCACCTTGGGCAATTAC 1644
Db 682 TATCAGCATCTACCGGTATTAACCTGCCCAAGAAATCATTTCCGGTTCACCTTGGGCAATTAC 741
Qy 1645 AGAATAGGAACAAAGTTGTGATTCAGAGGACAGACTACTAAAGCCAGGTTTTATATG 1704
Db 742 AGAATAGGAACAAAGTTGTGATTCAGAGGACAGACTACTAAAGCCAGGTTTTATATG 801
Qy 1705 GAGAGAACATGCCACCAAGATCAGAAATAGGTTACCGGAATTCATTGAGACAAACTA 1764
Db 802 GAGAGAACATGCCACCAAGATCAGAAATAGGTTACCGGAATTCATTGAGACAAACTA 861
Qy 1765 ACAAAACTAAACAGTCTATGCCATTTGGAAGAGTCCAGTTAACTTCTAAATAGCCAG 1824
Db 862 ACAAAACTAAACAGTCTATGCCATTTGGAAGAGTCCAGTTAACTTCTAAATAGCCAG 921
Qy 1825 ATTGTGATGTAAGACAGATGATTCAGTTGATGTCCTGTTTATGAAAGACAAACCTCTA 1884
Db 922 ATTGTGATGTAAGACAGATGATTCAGTTGATGTCCTGTTTATGAAAGACAAACCTCTA 981
Qy 1885 GATCAGAAATGCCAGATTTGGTTGCTGCTGATCTAAACCAAGTGGAAATGATTCCTGTG 1944
Db 982 GATCAGAAATGCCAGATTTGGTTGCTGCTGATCTAAACCAAGTGGAAATGATTCCTGTG 1041
Qy 1945 AATTAAAGAAATTTAAAGTCTGTTCAAAATAGTCAATTCAGAGAACCTCTGCTCAGATG 2004
Db 1042 AATTAAAGAAATTTAAAGTCTGTTCAAAATAGTCAATTCAGAGAACCTCTGCTCAGATG 1101
Qy 2005 AAAAGAGTTCTGAACCTTATTAATTAATCAATTAACCTGGAAGATTAACCGGAATCAA 2064
Db 1102 AAAAGAGTTCTGAACCTTATTAATTAATCAATTAACCTGGAAGATTAACCGGAATCAA 1161
Qy 2065 GTCTTCTAGCTAAATTAAGAAACTAAAGAGTATCAAGAACCCAGAGGTTCCAGAGATG 2124
Db 1162 GTCTTCTAGCTAAATTAAGAAACTAAAGAGTATCAAGAACCCAGAGGTTCCAGAGATG 1221
Qy 2125 ACCGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGTGTATTAACCAATCTCTGCTGCAT 2184
Db 1222 ACCGAAACAGTGGCAATCTAAGAGAAAGTCAGAGTGTATTAACCAATCTCTGCTGCAT 1281
Qy 2185 CTTCAAATCACTGGCAGATTCGGAGTTAGCCGAAAAGTTAATACAGAGCAGAAACATA 2244
Db 1282 CTTCAAATCACTGGCAGATTCGGAGTTAGCCGAAAAGTTAATACAGAGCAGAAACATA 1341
Qy 2245 CCACCTTTGACCAACCTGTCTTTTCACTTTCAAAACAGTCAACCAATATCAACATCTA 2304
Db 1342 CCACCTTTGACCAACCTGTCTTTTCACTTTCAAAACAGTCAACCAATATCAACATCTA 1401
Qy 2305 AATGGTTTGACCCAAAATCTATTTGTAAAGACACCAAGCAGCAATACCTTGGATGATTACA 2364
Db 1402 AATGGTTTGACCCAAAATCTATTTGTAAAGACACCAAGCAGCAATACCTTGGATGATTACA 1461
Qy 2365 TGAGCTGTTTGAACCTCAGTTGTAAGAAATGATCTTTCCACCTGCTTGTGTCAGTGTCAA 2424
Db 1462 TGAGCTGTTTGAACCTCAGTTGTAAGAAATGATCTTTCCACCTGCTTGTGTCAGTGTCAA 1521
Qy 2425 CACCTTTATGGCAACCTGCTGTTTCCAGAGCAGCAACAGATCAATATCTTGCACCTCCAC 2484
Db 1522 CACCTTTATGGCAACCTGCTGTTTCCAGAGCAGCAACAGATCAATATCTTGCACCTCCAC 1581

Qy 2485 TTCAAATTTTACAGGTTTTAGCATCTTCTTCAAGCAAAATGAATGCAATTTTCGGTTAAAGGAA 2544
Db 1582 TTCAAATTTTACAGGTTTTAGCATCTTCTTCAAGCAAAATGAATGCAATTTTCGGTTAAAGGAA 1641
Qy 2545 GAATTTATTTCCATATTTAAAGCAGATAGGAAGTGGAGGTTCAAGCAAGGTATTTTCAGGTGT 2604
Db 1642 GAATTTATTTCCATATTTAAAGCAGATAGGAAGTGGAGGTTCAAGCAAGGTATTTTCAGGTGT 1701
Qy 2605 TAAATGAAAGAAACAGATATATGCTATAAATAATGTGAACCTTGAAGAAGCAGATAACC 2664
Db 1702 TAAATGAAAGAAACAGATATATGCTATAAATAATGTGAACCTTGAAGAAGCAGATAACC 1761
Qy 2665 AAACCTCTTGATAGTTACCGGAAACGAAATAGCTTTTGAATAAACTACAACAACACAGTG 2724
Db 1762 AAACCTCTTGATAGTTACCGGAAACGAAATAGCTTTTGAATAAACTACAACAACACAGTG 1821
Qy 2725 ATAAGATCATCCGACTTTATGATTAATGAATCAACGACAGTACATCTACATGTTAATGG 2784
Db 1822 ATAAGATCATCCGACTTTATGATTAATGAATCAACGACAGTACATCTACATGTTAATGG 1881
Qy 2785 AGTGTGGAAATTTGATCTTAATAGTTGGCTTTAAAAAGAAAAATCCATTTGATCCATGGG 2844
Db 1882 AGTGTGGAAATTTGATCTTAATAGTTGGCTTTAAAAAGAAAAATCCATTTGATCCATGGG 1941
Qy 2845 AACGMAAGGTTACTCGGAAAAATATGTTAGAGGCAAGTTCACACAATCATCAACATGGCA 2904
Db 1942 AACGMAAGGTTACTCGGAAAAATATGTTAGAGGCAAGTTCACACAATCATCAACATGGCA 2001
Qy 2905 TTGTTTACAGTGTATCTTAAACCCAGTAACTTTCTGATAGTTGATGGAATGCTTAAAGCTAA 2964
Db 2002 TTGTTTACAGTGTATCTTAAACCCAGTAACTTTCTGATAGTTGATGGAATGCTTAAAGCTAA 2061
Qy 2965 TTGATTTTGGGATTCGAAACCAATGCAACCAAGTACACAACAGTGTGTTTAAAGATTCTC 3024
Db 2062 TTGATTTTGGGATTCGAAACCAATGCAACCAAGTACACAACAGTGTGTTTAAAGATTCTC 2121
Qy 3025 AGTGTGGCAAGTAAATTAATGTCACAGAGCAATCAAGAGATGCTTCTCCTCCAGAG 3084
Db 2122 AGTGTGGCAAGTAAATTAATGTCACAGAGCAATCAAGAGATGCTTCTCCTCCAGAG 2181
Qy 3085 AGAATGGGAAATCTAAGTCAAAAGATAAGCCCAAAAGTATGATTTTGGCTTTAGGATGTA 3144
Db 2182 AGAATGGGAAATCTAAGTCAAAAGATAAGCCCAAAAGTATGATTTTGGCTTTAGGATGTA 2241
Qy 3145 TTTTGTACTATATGATTTACGGGAAAAACCACTTTACAGAGATAATTAATCAAGATTCTA 3204
Db 2242 TTTTGTACTATATGATTTACGGGAAAAACCACTTTACAGAGATAATTAATCAAGATTCTA 2301
Qy 3205 AATTACATGCCATTAATGATCCTAATCATGAATTTGAATTTTCCGATATTCAGAGAAAG 3264
Db 2302 AATTACATGCCATTAATGATCCTAATCATGAATTTGAATTTTCCGATATTCAGAGAAAG 2361
Qy 3265 ATCTTCAAGATGTTTAAAGTGTGTTTAAAGAGGACCCAAAACAGAGGATATCCATTC 3324
Db 2362 ATCTTCAAGATGTTTAAAGTGTGTTTAAAGAGGACCCAAAACAGAGGATATCCATTC 2421
Qy 3325 CTGAGCTCTGCTCATCATATGTTCAAAATTCBAACCTCAGTTCACCAATAGGCGCA 3384
Db 2422 CTGAGCTCTGCTCATCATATGTTCAAAATTCBAACCTCAGTTCACCAATAGGCGCA 2481
Qy 3385 AGGGAACCACTGAAGAAATGAATATGTTTGGGCAACTTTGTTGGTCTGAATTTCTCCTA 3444
Db 2482 AGGGAACCACTGAAGAAATGAATATGTTTGGGCAACTTTGTTGGTCTGAATTTCTCCTA 2541
Qy 3445 ACTCCATTTTGAAGCTGCTAAACCTTTATATGAACATATAGTGGTGGTGAAGTCA 3504
Db 2542 ACTCCATTTTGAAGCTGCTAAACCTTTATATGAACATATAGTGGTGGTGAAGTCA 2601
Qy 3505 ATCTTTCATCTCCCAAGCTTTTGAAGAAAAAGGGGAAAAAATGATTTGAGTATTTC 3564
Db 2602 ATCTTTCATCTCCCAAGCTTTTGAAGAAAAAGGGGAAAAAATGATTTGAGTATTTC 2661

1321 CGAAAGTAAATACAGAGCAGAAATACCACTTTGAGCAACTGCTTTTCAGTTTCA 1380
2277 AAACAGTCACCAACCAATATCAACATCTAATGATGTTTGACCAAAATCTATTGTGAACA 2336
1381 AAACAGTCACCAACCAATATCAACATCTAATGATGTTTGACCAAAATCTATTGTGAACA 1440
2337 CCAAGCAGCAATACCTGGATGATTACATGAGCTGTTTGAAGACTCCAGTTGTAAGAAT 2396
1441 CCAAGCAGCAATACCTGGATGATTACATGAGCTGTTTGAAGACTCCAGTTGTAAGAAT 1500
2397 GACTTTCCACTGCTGTTGATGTTGACACCTTATGCGCAACCTGCTGTTTCCAGCAG 2456
1501 GACTTTCCACTGCTGTTGATGTTGACACCTTATGCGCAACCTGCTGTTTCCAGCAG 1560
2457 CAAGCAGTCAAAATACCTGACCTCCACTTCAAAATTTACAGGTTTACGATCTTCTTCA 2516
1561 CAAGCAGTCAAAATACCTGACCTCCACTTCAAAATTTACAGGTTTACGATCTTCTTCA 1620
2517 GCAAAATGAATGCAATTCGGTTTAAAGGAAGATTTTATCCATATTAAGCAGATAGGAAGT 2576
1621 GCAAAATGAATGCAATTCGGTTTAAAGGAAGATTTTATCCATATTAAGCAGATAGGAAGT 1680
2577 GGAGGTTCAAGCAAGGATTTTACGAGTTTAAATGAAAAGAAACAGATATATGCTATAAAA 2636
1681 GGAGGTTCAAGCAAGGATTTTACGAGTTTAAATGAAAAGAAACAGATATATGCTATAAAA 1740
2637 TATGTGAATTTAGAAAGCAGATTAACCAAACTCTTGATAGTTACCGGAACGAAATAGCT 2696
1741 TATGTGAATTTAGAAAGCAGATTAACCAAACTCTTGATAGTTACCGGAACGAAATAGCT 1800
2697 TATTGTAATAACTACAAACACAGTGAAGATCATCGACTTTATGATTTATGAAATC 2756
1801 TATTGTAATAACTACAAACACAGTGAAGATCATCGACTTTATGATTTATGAAATC 1860
2757 ACGGACGAGTACATCTACATGTTTAAAGGAGTGAAGATTTGATGTTTATGATGTTGCTT 2816
1861 ACGGACGAGTACATCTACATGTTTAAAGGAGTGAAGATTTGATGTTTATGATGTTGCTT 1920
2817 AAAAAGAAAAATCCATTTGATCCATGGGAACGCAAGAGTTTACTGGAAAAATATGTTAGAG 2876
1921 AAAAAGAAAAATCCATTTGATCCATGGGAACGCAAGAGTTTACTGGAAAAATATGTTAGAG 1980
2877 CGAGTTACACAATCCATCAATGCGATTTGTTACAGTGTATCTTAACACAGCTAACTTT 2936
1981 CGAGTTACACAATCCATCAATGCGATTTGTTACAGTGTATCTTAACACAGCTAACTTT 2040
2937 CTGATAGTTGATGGAATGCTAAAGCTAAATGATTTGGGATTGCAAAACCAATGCAACCA 2996
2041 CTGATAGTTGATGGAATGCTAAAGCTAAATGATTTGGGATTGCAAAACCAATGCAACCA 2100
2997 GATCAACAAGTGTGTTTAAAGATCTCAGGTTGGCAGTAAATTTATATGCCCACCAAGAA 3056
2101 GATCAACAAGTGTGTTTAAAGATCTCAGGTTGGCAGTAAATTTATATGCCCACCAAGAA 2160
3057 GCAATCAAGATATGTTCTCTCCAGAGAGATGGAATCTAAGTCAAGATGAAGCCCC 3116
2161 GCAATCAAGATATGTTCTCTCCAGAGAGATGGAATCTAAGTCAAGATGAAGCCCC 2220
3117 AAAAGTGATGTTGGTCTCTAGGATGTTTGTACTATATGACTTACGGGAAAACACCA 3176
2221 AAAAGTGATGTTGGTCTCTAGGATGTTTGTACTATATGACTTACGGGAAAACACCA 2280
3177 TTTGAGCAGATAAATTAATCAGATTTCTAATTAATGATGCGCAATTAATGATCAATGAA 3236
2281 TTTGAGCAGATAAATTAATCAGATTTCTAATTAATGATGCGCAATTAATGATCAATGAA 2340
3237 ATTGAATTTCCGATATTTCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAA 3296
2341 ATTGAATTTCCGATATTTCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAA 2400
3297 AGGAGCCCAAAACAGAGATATCCATCTCGAGCTCTCGGCTCATCCATATGTTCAAAT 3356

2401 AGGAGCCCAAAACAGAGGATATCCATCTCTGAGCTCTCTGGCTCATCATATGTTCAAAT 2460
3357 CAAAATCTCATCTAAGTTAAACCAATGCGCAAGGAAACCACTGAGAAATGAAATATGTTCTG 3416
2461 CAAAATCTCATCTAAGTTAAACCAATGCGCAAGGAAACCACTGAGAAATGAAATATGTTCTG 2520
3417 GGCACAACTGTTGGTCTGAATCTCTAATCTCTAATCTCTAATCTCTAATCTCTAATCTCTAAT 3476
2521 GGCACAACTGTTGGTCTGAATCTCTAATCTCTAATCTCTAATCTCTAATCTCTAATCTCTAAT 2580
3477 GAACTATAGTGGTGGTGAAGAGTCATTAATCTCTAATCTCTAATCTCTAATCTCTAATCTCTAAT 3536
2581 GAACTATAGTGGTGGTGAAGAGTCATTAATCTCTAATCTCTAATCTCTAATCTCTAATCTCTAAT 2640
3537 AGGGGAAAAAATGATTTGCACTTATGTAATGTCAGATGTCAGATGAGGTATATAATATATG 3596
2641 AGGGGAAAAAATGATTTGCACTTATGTAATGTCAGATGTCAGATGAGGTATATAATATATG 2700
3597 GACTGTTATACCTCTTGAATCCCTGTGGAAATCTAATTTGAAGACAAACATCACTCTGAAG 3656
2701 GACTGTTATACCTCTTGAATCCCTGTGGAAATCTAATTTGAAGACAAACATCACTCTGAAG 2760
3657 TGTATACGCAAAAAAATTCAGTGAGATTTCTTAAAGGAAACTGTAAAAATAGCA 3716
2761 TGTATACGCAAAAAAATTCAGT-AGATTATCTTTAAAGGAAACTGTAAAAATAGCA 2819
3717 CCACTTATGGCACTGATATATGTTAGACTGTTTCTCTGTTTATGCTCTTGTGTAAT 3776
2820 CCACTTATGGCACTGATATATGTTAGACTGTTTCTCTGTTTATGCTCTTGTGTAAT 2879
3777 CTACTTGACATCATTTTACTCTTGAATAGTGGTGGATAGCAAGTATATCTTAAAAAC 3836
2880 CTACTTGACATCATTTTACTCTTGAATAGTGGTGGATAGCAAGTATATCTTAAAAAC 2939
3837 TTGTGAAAAAAGTTTGTGGCTAAAAATGA 3866
2940 TTGTGAAAAAAGTTTGTGGCTAAAAATGA 2969

RESULT 7

AK095599 2419 bp mRNA linear PRI 15-JUL-2002
LOCUS AK095599
DEFINITION Homo sapiens cDNA FLJ38280 fis. clone FCBFF3005497, highly similar to DUAL SPECIFICITY PROTEIN KINASE TKK (EC 2.7.1.1-).
ACCESSION AK095599
VERSION AK095599.1 GI:21754889
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 Suzuki, O., Sasaki, N., Aotsuka, S., Shoji, T., Ichihara, T., Shiohata, N., Matsumoto, K., Hirano, M., Sano, S., Nomura, R., Yoshikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Momiya, H., Onogawa, S., Kaeriyama, S., Satoh, N., Matsunawa, H., Takahashi, E., Terasaki, R., Kuga, N., Kuroda, A., Satoh, I., Kamata, K., Takami, S., Sato, H., Oca, T., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Sato, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuko, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
REFERENCE 2 (bases 1 to 2419)
Isogai, T. and Yamamoto, J.
Direct Submission
TITLE
JOURNAL
AUTHORS
JOURNAL
COMMENT

Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5' - & 3' - end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

FEATURES

Location/Qualifiers
1..2419
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FCBFB3005497"
/tissue_type="brain"
/clone_lib="FCBFB3"
/dev_stage="fetal"
/note="cloning vector: pME189FL3"

ORIGIN

Query Match 47.8%; Score 1847.2; DB 9; Length 2419;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 28; Indels 15; Gaps 2;

Qy 887 CATTGTTGATGTGAATGTAAGCTGTACAGACTTCCTAGAAAACAGTTGGGTTCATCT 946
Db 481 CGTAGTATAGCGACGGGGAATTCAAACGTGTTGCGAAAGGAGTTTGGGTTCATCT 540

Qy 947 TTTCATTTCCCGAGTCAGTTTCTG-----TAGAATGGAAATCCGAGGATTT 994
Db 541 TTTCATTTCCCGAGTCAGTTTCTGAGTTTTTCTTAGAATGGAAATCCGAGGATTT 600

Qy 995 AAGTGGCAGAGATTGCAATGATTCATTAATGAACAAAGTGAGAGACATTAATAATAA 1054
Db 601 AAGTGGCAGAGATTGCAATGATTCATTAATGAACAAAGTGAGAGACATTAATAATAA 660

Qy 1055 GTTTAAATGAGACCTTACTGATGAACTAAGCTTGAATAAATTTCTGCTGATACTAC 1114
Db 661 GTTTAAATGAGACCTTACTGATGAACTAAGCTTGAATAAATTTCTGCTGATACTAC 720

Qy 1115 AGATAACTCGGGAAGCTTTAAACAAATTTATGATGATGGCAACAAACCCAGAGGACTGGTT 1174
Db 721 AGATAACTCGGGAAGCTTTAAACAAATTTATGATGATGGCAACAAACCCAGAGGACTGGTT 780

Qy 1175 GAGTTTGTGCTAAACTAGAGAAAACAGTGTTCGCTAAGTGATGCTCTTTTAAATAA 1234
Db 781 GAGTTTGTGCTAAACTAGAGAAAACAGTGTTCGCTAAGTGATGCTCTTTTAAATAA 840

Qy 1235 ATTGATTTGCTGTACAGTCAAGCAATTTGAAGCGCTTCCGCCAGATAAATATGGCCAAA 1294
Db 841 ATTGATTTGCTGTACAGTCAAGCAATTTGAAGCGCTTCCGCCAGATAAATATGGCCAAA 900

Qy 1295 TGAGAGTTTGTGTAGAAATCAAGTGAGATTGTGTGAATTAAGCTATTCAAGAGCCAGA 1354
Db 901 TGAGAGTTTGTGTAGAAATCAAGTGAGATTGTGTGAATTAAGCTATTCAAGAGCCAGA 960

Qy 1355 TGATGACGAGTACTTTCAATGGCCAGAGCAAACTGCAAGAAATTTGCTTTTGTCCA 1414
Db 961 TGATGACGAGTACTTTCAATGGCCAGAGCAAACTGCAAGAAATTTGCTTTTGTCCA 1020

Qy 1415 TATATCTTTTGCAAAATTTGAAGCTGCAAGGTAATGTCAAAAAGGTAACAACTTCT 1474
Db 1021 TATATCTTTTGCAAAATTTGAAGCTGCAAGGTAATGTCAAAAAGGTAACAACTTCT 1080

Qy 1475 TCAAAAGCTGTAGAACTGGAGAGTACCAGTATAGAAATGCTGGAATTTGCCCTGCGGAA 1534
Db 1081 TCAAAAGCTGTAGAACTGGAGAGTACCAGTATAGAAATGCTGGAATTTGCCCTGCGGAA 1140

Qy 1535 TTTAAACCTCCAAAAGAGCAGCTGCTTTTCAGAGGAGGAAAAGAAATTTATCAGCATC 1594
Db 1141 TTTAAACCTCCAAAAGAGCAGCTGCTTTTCAGAGGAGGAAAAGAAATTTATCAGCATC 1200

Qy 1595 TAGGTTATTAAGTCCCAAGAAATCATTTTCCGTTTCACTTGGCATTTCAGAAATAGGAA 1654
Db 1141 TAGGTTATTAAGTCCCAAGAAATCATTTTCCGTTTCACTTGGCATTTCAGAAATAGGAA 1654

Db 1201 TACGGTATTAACTGCCCAAGAAATCATTTTCGGTTCACTTGGGCATTTACAGATAGGAA 1260
Qy 1655 CAACAGTTGTGATTCAGAGAGACAGACTAATAAGCCAGAGTTTATATATGAGAGAACAT 1714
Db 1261 CAACAGTTGTGATTCAGAGAGACAGACTAATAAGCCAGAGTTTATATATGAGAGAACAT 1320
Qy 1715 GCCACCAAGATGAGAGAAATAGGTTACCGGAATTCATTGAGACAAACTACAAATATA 1774
Db 1321 GCCACCAAGATGAGAGAAATAGGTTACCGGAATTCATTGAGACAAACTACAAATATA 1380
Qy 1775 ACAGTCATGCCCATTTGGAAGAGTCCAGTCTTAACCTTAAATAGCCAGATTTGATGT 1834
Db 1381 ACAGTCATGCCCATTTGGAAGAGTCCAGTCTTAACCTTAAATAGCCAGATTTGATGT 1440
Qy 1835 GAAGACAGATGATTCAGTTGTACCTTGTATGAAAAGACAAACCTCTAGATCAGAAATG 1894
Db 1441 GAAGACAGATGATTCAGTTGTACCTTGTATGAAAAGACAAACCTCTAGATCAGAAATG 1500
Qy 1895 CCGAGATTTGGTTGTGCTGATCTAAACCAAGTGGAAATGATTCCTGTGAAATTAAGAAA 1954
Db 1501 CCGAGATTTGGTTGTGCTGATCTAAACCAAGTGGAAATGATTCCTGTGAAATTAAGAAA 1560
Qy 1955 TTTAAAGTCTGTTCAAATAAGTCAATTTCAAGGAACCTCTGCTGTCAGATGAAAGAGTTC 2014
Db 1561 TTTAAAGTCTGTTCAAATAAGTCAATTTCAAGGAACCTCTGCTGTCAGATGAAAGAGTTC 1620
Qy 2015 TGAATCTTATTATGATTCATTAACCTTGAAGAAATAAAACGGAATCAAGTCTTCTAGC 2074
Db 1621 TGAATCTTATTATGATTCATTAACCTTGAAGAAATAAAACGGAATCAAGTCTTCTAGC 1680
Qy 2075 TAAATTAGAAGAACTAAGAGTATCAAGAACAGAGGTTCCAGAGAGTAAACGAGAAACA 2134
Db 1681 TAAATTAGAAGAACTAAGAGTATCAAGAACAGAGGTTCCAGAGAGTAAACGAGAAACA 1740
Qy 2135 GTGCAAGCTAAGAGAAAGTCAGAGTGTATTAAACCAAGTCCCTGCTGCATCTTCAAAATCA 2194
Db 1741 GTGCAATCTAAGAGAAAGTCAGAGTGTATTAAACCAAGTCCCTGCTGCATCTTCAAAATCA 1800
Qy 2195 CTGCAAGTTCGGAGTTAGCCCGAAAGTTAATACAGAGCAGAGAAACATACCACTTTGA 2254
Db 1801 CTGCAAGTTCGGAGTTAGCCCGAAAGTTAATACAGAGTTCCTGCTGCATCTTCAAAATCA 1857
Qy 2255 GCAACCTGCTTTTTCAGTTTCAAAACAGTCAACCAAGTATCACTCACTCAATGTTTGA 2314
Db 1858 GCAACCTGCTTTTTCAGTTTCAAAACAGTCAACCAAGTATCACTCACTCAATGTTTGA 1917
Qy 2315 CCCAAATCTATTGTTAAGACACCAAGCAGCAATACCTTGGATGATTACATGAGCTGTTT 2374
Db 1918 CCCAAATCTATTGTTAAGACACCAAGCAGCAATACCTTGGATGATTACATGAGCTGTTT 1977
Qy 2375 TAGAATCTCAGTTGTAAGAAATGATTTCCACCTGCTTGTGTCAGTTGTCACACCTTAGG 2434
Db 1978 TAGAATCTCAGTTGTAAGAAATGATTTCCACCTGCTTGTGTCAGTTGTCACACCTTAGG 2037
Qy 2435 CCAACCTGCTGTTTCCAGCAGCAACAGCATCAATACCTTGGCACTCCACTTCAAAATTT 2494
Db 2038 CCAACCTGCTGTTTCCAGCAGCAACAGCATCAATACCTTGGCACTCCACTTCAAAATTT 2097
Qy 2495 ACAGTTTTCAGCATCTTCTTCAGCAAAATGAATGCAATTCGGTTAAAGGAGAAATTTATTC 2554
Db 2098 ACAGTTTTCAGCATCTTCTTCAGCAAAATGAATGCAATTCGGTTAAAGGAGAAATTTATTC 2157
Qy 2555 CATATTAAAGCAGATAGGAAGTGGAGGTTCAAGCAAGGTAATTTTCAGGTGTTAAATGAAA 2614
Db 2158 CATATTAAAGCAGATAGGAAGTGGAGGTTCAAGCAAGGTAATTTTCAGGTGTTAAATGAAA 2217
Qy 2615 GAAACAGATATATGCTATAAATAATGTGAATTTAGAAAGAGCAGATTAACCAACTCTTGA 2674
Db 2218 GAAACAGATATATGCTATAAATAATGTGAATTTAGAAAGAGCAGATTAACCAACTCTTGA 2277
Qy 2675 TAGTTACCGGACCAAGATAGCTTATTTGATTAACATACACACAGCAGTTCATAGATCAT 2734
Db 2278 TAGTTACCGGACCAAGATAGCTTATTTGATTAACATACACACAGCAGTTCATAGATCAT 2337

Qy 2735 CCGACTTTATGATTATGAATACGGACAGTACATCTACATGATGATGAGTGTGAAA 2794
 |||||
 Db 2338 CCGACTTTATGATTATGAATACGGACAGTACATCTACATGATGATGAGTGTGAAA 2397
 |||||

Qy 2795 TATTGATCTTAATAGTTGGCTT 2816
 |||||
 Db 2398 TATTGATCTTAATAGTTGGCTT 2419
 |||||

RESULT 8
 BC058851 2906 bp mRNA linear ROD 21-OCT-2003
 LOCUS Mus musculus Ttk protein kinase, mRNA (cdna clone MGC:60685
 IMAGE:30023533), complete cds.
 ACCESSION BC058851
 VERSION BC058851.1 GI:37589052
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2906)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, E.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 23388257
 12477932
 2 (bases 1 to 2906)
 Strausberg, R.
 Direct Submission
 Submitted (01-OCT-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Janet Rossant and Tilo Kunath (Samuel
 Lunenfeld Research Institute, Canada)
 cDNA Library Preparation: Yulan Piao and Minoru Ko (National
 Institute on Aging, NIH: <http://igsun.grc.nia.nih.gov/cDNA/>)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegue, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LILNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 111 Row: m Column: 18

FEATURES
 Source

This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6678462.
 Location/Qualifiers
 1..2906
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="BS/EGFP transgenic ICR mice"
 /db_xref="taxon:10090"
 /clone="MGC:60685 IMAGE:30023533"
 /tissue_type="Trophoblast Stem Cell, mouse, 3.5 dpc"
 /clone_lib="NIA Mouse Trophoblast Stem Cell cDNA Library
 (Long)"
 /lab_host="DH10B"
 /note="Vector: pSPOR11"
 1..2906
 /gene="Ttk"
 /note="synonyms: PYT, esk, Esk1"
 /db_xref="LocusID:22137"
 /db_xref="MGI:1194921"
 76..2568
 /codon_start=1
 /product="Ttk protein"
 /protein_id="AAH58851.1"
 /db_xref="GI:37589053"
 /db_xref="LocusID:22137"
 /translation="MEAEELIGSSVTIDISIMSKMDIKNKINEDCTDELSIKICADH
 TETVNCIRVGNTPENWLNFLKSKNSPLNDLLNKLIGYSAIENVLPDKYQSN
 EBFARIQVLAELKAIQFDDARDIFQWARENCKAFVHVHVSFAQFELSQGNLKSEQ
 LHKAVETGAVGLQWETAMRNHLQKKQLLPEDDKSVASATVLSAGEPPFSSUGNV
 QNRSSCSGRGQAGLAVLYGENLPQDAEVRHONPQKTHAAKSCPPGRVPLNLLN
 SPDVYKTDSSAVTQITTKGSDPDRDAILPGSRPRGSDSVELRGKIQITLYLKDLSL
 VNEKSSLSMDLIALKKSTDSLTLEETKEIAERRPMOWSTRKPCVQPNDAAF
 ALRHVPVTLKADKESPIVVKWLDPKSACETPSSSLDDYMKCFKPVVKNDPPP
 ACPSTPQSLARLQOQQOQGLSTFQSLQISGSSINECISVNGRIYSLIKLIDSGG
 SSKVFLVNEKQINAIYNEIAFLNPKSACETPSSSLDDYMKCFKPVVKNDPPP
 TQYIYVMMECNIDLNWKKKKSINPWKSYWKNLMEAVHIIHQHGIHSDLKPA
 NFVIVDGLKIDFGIANGQPDITSIVKDSQVGTNYNPAEIAIDMSRRNSKIRT
 KVSPPSDVMSGLCIVYMTYTPFQHIINOVSKLHAIINPAHEIEFPEISKDLRDV
 LKCCVLRPKESIS:PELLTHPYVOIOPHPGSGMARGATDEMKYVLQGLVGLNSPNSI
 LKTAHLRYNCGEGQSSSSSKTFDKKRRK"

misc_feature

1567..2367
 /note="S_TK; Region: Serine/Threonine protein kinases,
 catalytic domain"
 /db_xref="CDD:smart00220"

ORIGIN

Query Match 42.3%; Score 1634.6; DB 10; Length 2906;
 Best Local Similarity 76.8%; Pred. No. 0;
 Matches 2193; Conservative 0; Mismatches 559; Indels 103; Gaps 12;

Qy 971 TGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAAATGCAAAATTGATTCCTAATGAA 1030
 |||||

Db 69 TTTAGAAATGGAGCTGAAGAGTTAATTCGCAGCAGTGTGACGATTTGATTCATCATGAG 128
 |||||

Qy 1031 CAAAGTGAGAGACATTAATAAAGTTTAAATGAAGACCTTACTGATGAACCTAGCTT 1090
 |||||

Db 129 CAAATGAGAGATATTAATAAAG---ATAAATGAAGACTGTACTGATGAGCTAGCTT 185
 |||||

Qy 1091 GATATAAATTTCTGTGATACAGATAAATCGGGAACCTGTTAAACCAATTATGATGAT 1150
 |||||

Db 186 GTCTAAATCTGTG-----CCGATCACCGGAACTGTTAACCAATTATGAGGCT 236
 |||||

Qy 1151 GCGAAACAAACCCAGAGACTGGTTGAGTTGTTGCTCAAACTAGAGAAAAACAGTGTTC 1210
 |||||

Db 237 TGGGAACACCCAGAGAACTGGTTGAAATTTCTTGTAACACTAGAGAAAAACAGCTCACC 296
 |||||

Qy 1211 GCTAAGTGATGCTCTTTTAAATAAATTTGATCGTTCAGTCAAGCAATTGAACGGCT 1270
 |||||

Db 297 TCTAAATGACGATCTTTTAAATAAGCTGATTTGTCGCTATAGTCAACGCGATTGAAGTACT 356
 |||||

Qy 1271 TCCCCCAGATAAATATGGCCAAATGAGAGTTTGTGTAGAAATCAAGTGAGATTTCTGTA 1330
 |||||

Db 357 TCCTCCAGATAAATACGCCAGAAATGAGAGCTTTGCTCGAATACAAAGTGAGACTTGTCTGA 416
 |||||

1331 QY ATTTAAAGCTATTCAAGAGCCAGATGATGCACTGACTACTTTCAATATGGCCAGAGCAAA 1390
1417 Db ACTAAAGCTATTCAAGAGCCAGATGATGCGCGTGACTACTTCCAGATGGCCAGGAGAAA 476
1391 QY CTGCAAGAAATTCCTCTTTGTTTCATATATCTTTTGGCAAAATTTGAAGTGTCAAGGTAA 1450
1477 Db CTGCAAGAAATTCCTCTTTGTTTCATATATCTTTTGGCAAAATTTGAAGTGTCAAGGTAA 536
1451 QY TGTCAAAAAAGTAAACAATCTTTCAAAAAAGCTGTAGAAACGTGGAGCAGTACCACTAGA 1510
1537 Db TCTTAAAAAAGTGAAGCAGCTTCTTCATAAAGCTGTAGAGACTGGGGCGGTGCGCTGCA 596
1511 QY AATGCTGGAATTCCTCTGGGAATTTAAACCTCCAAAAAAGAGAGCTGCTTTCAAGGA 1570
1597 Db GATGCTGGAGACGCCATGGTAACTTACACCTCCAAAAAAGAGAGCTGCTTTCCGAGGA 656
1571 QY GGAAGAAGAAATTTATCAGCATCTACGGTATTAACTGCCCCAAGAAATCAATTTCCGGTTC 1630
1657 Db GGACAGAGAGTGTCTCAGCATCGACAGTACTAAAGTGGCCAGAGCGGTCTTCCAGCTC 716
1631 QY ACTTGGGCATTTACAGAAATAGGAACAAAGTGTGTGATTCAGAGGACAGACTACTAAAGC 1690
1717 Db ACTTGGAAATGTACAGAAATCGGAGCATCAGCTGTGTAGTCCAGAGGACAGGCTGGGGCAGC 776
1691 QY CAGGTTTTTATATGGAGAGCAATGCCACCAAGATGCAAGAAATAGGTTACCGAATTC 1750
1777 Db CAGGTTTTTATATGGAGAGAACTTGGCTCCCAAGATGCCAAAGTGAAGCATCAAAACCC 836
1751 QY ATTGAGACAAACTAAACAACTAAACAGTCAATGCCATTTGGAGAGTCCCACTTAACCT 1810
1837 Db CTTCAAGCAGACTCACGAGCTAAACGGTCTATGCCCTTTGGAAGAGTCCCACTCAATCT 896
1811 QY TCTAAATAGCCCATTTGATGTGAAGAGACAGATGATTCAGTTGTACCTTGTATTATGAA 1870
1897 Db TCTAAACAGCCCATTTCTATGTGAAGACAGATGATTCAGTTGTACCTTGTATTATGAA 956
1871 QY AAGACAAACCTCTAGATCAGAAATGCCAGATTTGGTTGTGCTTGGATCTAAACCAAGTGG 1930
1957 Db AAGAAAGGGTGGGACCCAGA---CCGAGACGGATCTGCGCGCTCCAGACCACTGG 1013
1931 QY AATGATCTCTGGAATTAAGAAATTTAAAGTGTGTCTTCAAAATAGTCAATTTCAAGAAC 1990
2014 Db CAGTGTATCTATGAATGAGAGGTTTAAAGCCCATTTCAAACTATCTATTTGAAAGACTC 1073
1991 QY TCTGGTGTGATGAAAGAGTTCTGAACCTATTATTTACTGATTTCAATTAACCTGAAGAA 2050
2074 Db TTTGGTGTCCAATGAAGAGTTCTGNAC---TTATGTCTGATTTAATAGCTTTGAAGAG 1130
2051 QY TAAAAAGGAATCAAGTCTTCTAGCTAAATTAAGAGAACTTAAAGATATCAAGAACCCAGA 2110
2131 Db TAAAAACAGATTCAAG---TCTAAACAAATTTGGAAGAACTTAAG-----CCAGA 1175
2111 QY GTTCCAGAGAGTAAACAGAAACAGTGGCAAGCTTAAGAGAAAGTCAAGTGTATTAAACCA 2170
2176 Db GATTGAGAAAGAGGCCCATGAGTGGCAGTCTACAGAAAGCCGAGTGTGTGTTCCA 1235
2171 QY GAATCTGTGCTCTTCTTCAAACTCACTGGCAGATTCGGGAGTTAGCCCGGAAAAAGTTAAATAC 2230
2236 Db GAACCTGTGCTCTTGTGACCCCTGGGACAGTTTCCAGATGTCACTCTGAAGGCTGACAA 1295
2231 QY AGAGCAGAAACATACCACTTTTGAAGCAACCTGTCTTTTTCAGTTTCAAAACAGTCAACCC 2290
2296 Db AG-----AGTCCACCACT 1307
2291 QY AATATCAACATCTAAATGTTTGAACCCAAAATCTATTGTGAAGACCAACAGCAGCAATAC 2350
2308 Db AATATCAGTTCTTAATGGCTTGATCCAAAGTCTGCTTGTGAGACCACTAGTAGAGCTC 1367
2351 QY CTTGGATGATATGATGAGCTGTTTTAGAACTCCAGTTGTAAAGATGATTTCCACCTGTC 2410
2368 Db CTTGGATGATATGATGAAATGTTTTAAGACTCCAGTTGTAAAGATGATTTTCCACCTGTC 1427

2411 QY TTGTCAAGTTGTCAACACCTTTATGGCCAACTGCTGTTTCCAGCAGCAACAGCATCAAT 2470
2428 Db CTGTCCATCATCAACACCTTTACAGCCAGCTTTCGCCGCTCCAGCAGCAACAGCAGCAGG 1487
2471 QY ACTTCCCACTCCACTTCAAAATTTACAGTTTTCAGCTTCTTTCAGCAAAATGATGAT 2530
2488 Db ACTCAGCACTCTCTCTTCAAGCTTTCAGATTTTCAGGTTCTTTCATCAATAAATGATGAT 1547
2531 QY TTTCGTTTAAAGGAGAAATTTTATTCATATTAAGCAGATAGCAAGTGGAGGTTTCAAGCAA 2590
2548 Db TTCAGTTAACGAGAAATTTTATTCATATTAAGCAGATAGCAAGTGGAGGTTTCCAGTAA 1607
2591 QY GGTATTTTCAGGTTTAAATGAAAGAAACAGATATATGCTATAAATATGTAACCTTAGA 2650
2608 Db GGTGTTTTCAGGTTTAAATGAAAGAAACAGATATATGCTATAAATATGTAACCTTAGA 1667
2651 QY AAGACAGATTAACCAACTCTTGTATGTTACCGAACGAAATAGCTTATTTGAATAAAT 2710
2668 Db AGACCCGATAGCCAAACTATTTGAGAGCTACCGCAACGAGATAGGTTTTCAGCAAACT 1727
2711 QY ACAACACAGATGATAGATCATCCGACTTTTATGATATGAAATCAGGACCACTAGTACAT 2770
2728 Db ACAGCAACAGATGATAGATCATCCGCTCTATGATTTGAAATCAGGAGCAGTACAT 1787
2771 QY CTACATGTTTAAAGTGTGAGAAATATTTGATCTTAAATAGTTGGCTTAAAGAAATAATC 2830
2788 Db CTACATGTTTAAAGTGTGAGAAATATTTGATCTTAAATAGTTGGCTTAAAGAAATAATC 1847
2831 QY CATTTGATCCATGGAAACGCAAGAGTTTACTGGAAAAATATTTGAGAGCAGTTTCAACAAT 2890
2848 Db CATCAATCCATGGAAACGCAAGAGCTTACTGGAAAAACATGTTGGAGCAGTACACATAAT 1907
2891 QY CCATCAACATGCGCATTTGTTACAGTGTCTTAAACAGCTAACTTTCTGATAGTTGAG 2950
2908 Db CCATCAGATGTTTGTTCATAGTGTCTTGAAGCTTGTGATAGTTGAGTTGAGTTG 1967
2951 QY AATGCTTAAAGCTTAAATTTGGGATTTGAAACCAATGCAACCAATGCAACCAAGTGT 3010
2968 Db AATGCTTAAAGCTTAAATTTGGGATTTGAAACCAATGCAACCAATGCAACCAAGTGT 2027
3011 QY TGTAAAGATTTCTCAGGTTGGCAGTAAATATATATGCCACCAAGCAATCAAGATAT 3070
2028 Db TGTAAAGATTTCTCAGGTTGGCAGTAAATATATATGCCACCAAGCAATCAAGATAT 2087
3071 QY GTCTTCTCCAGAGAGAAATGGAATCTAAGTCAAGATTAAGCCCAAGAGTGTGTTG 3130
2088 Db GTCTTCTCCAGAGAGAAATGGAATCTAAGTCAAGATTAAGCCCAAGAGTGTGTTG 2147
3131 QY GTCTTCTCCAGAGAGAAATGGAATCTAAGTCAAGTGTGTTTAAAGTGTGTTTAAAGAGGACCCAAACA 3190
2148 Db GTCTTCTCCAGAGAGAAATGGAATCTAAGTCAAGTGTGTTTAAAGTGTGTTTAAAGAGGACCCAAACA 2267
3191 QY TAATCAGATTTCTAAATTAAGTCAATGATGATCTTAAATCAATGATGATGATGATGATGATGAT 3250
2208 Db CAATCAGGTTCTTAAATTAAGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2267
3251 QY TAATCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAGAGGACCCAAACA 3310
2268 Db GATTTCCGAAAAAGATCTTCCAGAGCTGTCTTAAAGTGTGTTTAAAGAGGACCCAAACA 2327
3311 QY GAGATATCAATCTTCCAGCTCTGCTCATCATATGATGATGATGATGATGATGATGATGATGAT 3370
2328 Db GAGATATCAATCTTCCAGCTCTGCTCATCATATGATGATGATGATGATGATGATGATGATGAT 2387
3371 QY TAAACCAATGGCCAAAGGAAACCACTGAAGAAATGAAATATGTTTCTGGGCAACTTCTGTTG 3430
2388 Db CAGCCAAATGGCTAGGGAGCCACTGATGAAATGAAATATGTTGTTGGTCACTTCTGTTG 2447
3431 QY TCTGAATTTCTTAACTTCTGAAAGTGTCTTAAAGTGTCTTAAAGTGTCTTAAAGTGTCTTAAAGTGT 3490
2448 Db TCTGAATTTCTTAACTTCTGAAAGTGTCTTAAAGTGTCTTAAAGTGTCTTAAAGTGTCTTAAAGTGT 2507
3491 QY TGGTGAAGTCAATAATTTCTTCACTCTCCAGACTTTTGAAGAAAAAGGAAAAAATG 3550

```

Db      2508  TGTGAGAGTCAAGATCTTCGTCATCCAGACTTTTGACAAAGAGAGAGAAAGATG 2567
Qy      3551  ATTTCGAGTTATTCGTAATGTCAGATAGGAGGTATATAATATATATGAGCTGTATCTCT 3610
Db      2568  ATGCACAGCTACGTACAAACCAAGA-----ACACAGATTTGTTTCTCTGCGCATCTCT 2621
Qy      3611  TGAATCCCTGTGAATCTA-CAATTTGAAGACACATCACTCTGAAGTGT----- 3660
Db      2622  TGAATCTCTGAGAAATCTACCAAGTTGGAACCAACTCACCTGGATTTTATCAGTTAA 2681
Qy      3661  --ATCAGCAAAAAAATTCAGTGAGATTTATCTTTTAAAGAAAACTGTAATAATAGCAACC 3718
Db      2682  AACCAACAAACAAACTTCAGTAGATTAATCTCAAAAGAAAGCTGTAAGTT---AACC 2738
Qy      3719  ACTTATGGCACTGTATATATGAGCTGTGTTCT-CTGTTTATGCTCTGTGTAATC 3777
Db      2739  ACTCATAGCACTGTATATTAATATAGATTTGCTTTCTTTTATGCTTTTCTGTA 2798
Qy      3778  TACTTGACATCATTTACTCTTGAATAGTGGGTG 3812
Db      2799  AATCTGCTAATGTTTACGTTTGGACAGTGAATG 2833

RESULT 9
AX775731
LOCUS      AX775731      2860 bp      mRNA      linear      PAT 14-JUL-2003
DEFINITION Sequence 1 from Patent WO03048202.
ACCESSION AX775731
VERSION    AX775731.1 GI:32693449
KEYWORDS
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
  1 Matsuda, A. and Muramatsu, S.
  NF-kB activating gene
  Patent: WO 0304820-A 1 12-JUN-2003;
  Asahi Kasei Kabushiki Kaisha (JP)

FEATURES
  source            1..2860
                    /organism="Mus musculus"
                    /mol_type="mRNA"
                    /db_xref="taxon:10090"
                    56..2548
                    /note="unnamed protein product"
                    /codon_start=1
                    /protein_id="CAE11565.1"
                    /db_xref="GI:32693450"
                    /translation="MEAEELIGSSVTIDISMSKRDINKINKEDTDELSLKICADH
                    TETVQIMRVGNTPENLNLFLKLEKNSPLNDLLNKLIGRYSQAIEVLPPDKYQGN
                    ESFARIQVRLAELKAIQEPDDARDYFQWAKENCKFAFVHVHVSFAQFELSQNLKQSEQ
                    LLKRAVETGAVPLQMLTAMENLHLQKLLPEDKKSVASTVLSAQEPFSSSLGNV
                    QNSISCSBSQAGAAVLVGENLPQDAEVHONPFKOTHAARKSCFPQGVVNLIN
                    SPDYVTKDSAVTQLTRKSGSPDRDAIILPGSPGSDSYELRGLKPIOTIYLKQSL
                    VSNKSSLMEDLLALKSTUSLLKLEETKPEIAERRPMQWSTKPECVFQNPAP
                    APLRHPDVPYADKESPPISVPKWLDPKSACTPPSSSLDDYNNKCFKTPVKNRDPFP
                    ACPSPSTYSQLARLQOQOQGLSTPLQSLQISGSSINECISVNGRIYSILKQISGG
                    SSKYFQVINEKQINAIKYVNLKEDASQTIYSYRNEIAFLNKLQHSQDKIIRLYDVEI
                    TEQVLYMVECGNIDLSNKKKKSINPWERKSYKMNLEAVHIIHQHGVHSDLPKA
                    NFVLDWGLKLDPIGIANOMPDITSIVKDSQVGTVMYMAPEAIRDMSSRSKRT
                    KVSFPRVWSLGCILYNTYGTPTFQHLINQVSKLHAIINPAHEIEFPEISEKLDKRV
                    LKCLVRNPKRISIPPELLTHPYVQIQHPGSGQWARGATDEMKVILQVLGLSPNSI
                    LKTKTLIERYNCGEGQSSSSKTFDKKRKK"

ORIGIN
  Query Match      42.2%; Score 1633; DB 6; Length 2860;
  Best Local Similarity 76.8%; Pred No. 0;
  Matches 2192; Conservative 0; Mismatches 560; Indels 103; Gaps 12;

Qy      971  TGTAGAAATGGAAATCCGAGGATTTAAGTGGCAGAGAAATGCAATTTGATTCATAATGAA 1030

```


CDS	/cell_type="stem cell"		1631	ACTTGGGCATTTACAGAAATAGAAACACAGTGTGTATTCAGAGACAGACATCTATAAGC	1690
	38..2608		Db		
	/note="putative"		679	ACTTGGAAATGTACAGAAATAGAGCATCAGCTGTGAGTCCAGAGACAGGCTGGGGCAGC	738
	/codon_start=1				
	/product="esk kinase"		1691	CAGGTTTTATATGGAGAGAAATGCCACACACAGATGACAGAAATAGTTTACCGGAATTC	1750
	/protein_id="AAA37578.1"		Qy		
	/db_xref="GI:193110"		739	CAGGTTTTATATGGAGAGAAATGCCCTCCACAGATGCCGAAGTGGGCAATCAAAACCC	798
	/translations="MEAEELIGSVTIDSIMSKORDINKINEDCTDELSSKICADH				
	TEVNVQIMRVGNTPEWNLAFLLKLNSSPLNDLLNKLIGVSOAIEVLDPDKYGN		1751	ATTGAGACAAACTAACAAAATAAAACAGTCATGCCCATTTTGGAGAGTCCCGATTAACCT	1810
	ESPARQVRIAEUKAQEPDDARDYQMARENCCKPAPVHVSFAQPELSSGNLKSEBQ		Qy		
	LLHKAIVETGAPVQMLETAMRNHLQKQLLPEDBKSSVASTVLSAQPFPSSSUGNV		799	CTTCAAGCAGACTCAGCGAGCTAAACGGTCTATGCCCTCATGCCCTTGGAGAGTCCCGATCAATCT	858
	QNRISCEBRSQAVARVLYGENLPQDAEVRHQNPFOKTHAAKSCGPFPRFVNLLN		Db		
	SPDFYVKTSSAVTQTLRLASSVLPVYTCLLHLQALLALAGKSGSPDRDAILPG		1811	TCTAAATAGCCAGATTTGTATGTGAGACAGATGATTCAGTTGTACCTTGTATTATCAA	1870
	SRPFGDSVYELRGLKTIQTLKDSLVNEKSELSMSDLIALKSKTSSLSLTLEETKP		Qy		
	ELARERPMQCRPEKPCVQFQNPAPARLPVDPVTKADKESPPISVPEKLDPKXAS		859	TCTAAACAGCCAGATTTCTATGTGAAGACAGATAGCTCAGCTGTGACACAGTTAAACAAC	918
	ETSSSLDDYMKCFKTPVVKNDPFPACPSSTEPSOLASLQOQOQGLSTPLSLQIS		Db		
	GSSINECIRGIRISVILKQIGSGSKVQVNLNKKQINAIKYNLEADASQIES		1871	RAGACAAACCTCTAGATCAGAATG-----	1894
	YRNEIAPLNKLOQHSKIRLYDIETEQIYIMVMECNIIDLSNKKKKSINPWRKQ		Qy		
	SYWNKLEAVHIHQHGVHSIDLPANFIVDMGLKLIIDFGANQOPDTTISVCKDSQ		919	RAGGCTAGCCITTAAGCTCTGTACCTTGCCTCGTAAACCTGCCTCTGCATACAGCT	978
	VGTWVNYMAPAIRDMSSRENSKIRTKVSPRWSGLGILYMTYGRTPFOHINQV		Db		
	SKLHIINPAHEIEFPEISBKLDRLVKCLVRNPKRISIPPELLPHVVOIQPHGFS		1895	-----CCGAGATTTGGTTGCTCTGG	1915
	QMARGATDENKYLVLQVLGLNSPNSILKTAKTILYERYNCEGQDSSSSKTFDKRERK		Qy		
			979	GCTGGCGCTCGCAGGCTTGGCAAAAGGGCTCAGGACCCAGACGAGCGGATTTGCCCCGG	1038
			Db		
ORIGIN	Query Match 40.4%; Score 1562.4; DB 10; Length 2924;		1916	ATCTAAACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAATAG	1975
	Best Local Similarity 74.9%; Pred. No. 0;		Qy		
	Matches 2194; Conservative 0; Mismatches 561; Indels 175; Gaps 12;		1039	CTCCAGACACAGTGGCAGTGATTCCTATGAACCTGAGAGGTTTAAAGCCCATTCAAACTAT	1098
			Db		
			1976	TCATTTCAAGGAACCTCTGGTGCAGATGAAAAGAGTTCTGAACCTTATTTACTGATTC	2035
			Qy		
			1099	CTATTTGAAAGACTCTTTGGTGTCCAATGAAAAGAGTTCTGAAC---TTATGCTGTGATTT	1155
			Db		
			2036	AATAACCTGGAAGATAAACAAGAACTCAAGTCTTCTAGCTAAATTAGAGAACTAAAGA	2095
			Qy		
			1156	AATAGCTTGAAGAGTAAACAGATTTCAAG---TCTAACAAATTTGGAAGAAACTA----	1208
			Db		
			2096	GATCAAGAACCCAGAGGTTCCAGAGAGTAAACAGAAACAGTGGCAGCTAAGAGAAAGTC	2155
			Qy		
			1209	-----AGCCAGAGATTGCAAGAAAGAGCCCATGCACTGCGAGTCTACCCAGAAAGCC	1260
			Db		
			2156	AGAGTGTATTAACCGAATCTCTGCTGCATCTTCAATCACTCGCAGATTCGGAGTTAGC	2215
			Qy		
			1261	CGAGTGTGTGTTCCAGAACCTCTGCTGCTTTCACCCCTGCGGCGAGTTCAGATGTAC	1320
			Db		
			2216	CGGAAAGTTAATACAGAGCAGAAACATACCACTTTTGAGCAACCTGCTTTTCAGTTTC	2275
			Qy		
			1321	CCGAAAGGCTGA-----	1332
			Db		
			2276	AAAAAGTCAACCAATATCAACATCTAAATGGTTGACCCAAATCTATTTTGAAGAC	2335
			Qy		
			1333	CAAGAGTCAACCAATATCAAGTCTTAAATGGCTTGATCCAAAGTCTGCTTGTGAGAC	1392
			Db		
			2336	ACCAAGCAGAAATACCTTGGATGATTACATGAGCTGTTTGAACCTCCAGTTGTAAGAA	2395
			Qy		
			1393	ACCTAGTAGAGTCTCTTGGATGATTACATGAAATGTTTAAAGCTCCAGTTGTAAGAA	1452
			Db		
			2396	TGACTTTTCCACCTGCTTGTGAGTGTCAACACCTTATGGCCAACTGCTGTTTCAGCA	2455
			Qy		
			1453	TGACTTTTCCACCTGCTGCTTCCATCATCAACCTTACAGCGAGTTGCCCGCTCCAGCA	1512
			Db		
			2456	GCACAGAGTCAAAATCTTGGCCTCCACTTCAAAATTTTACAGTTTACAGCTCTCTTC	2515
			Qy		
			1513	GCAACAGCAGCGGAGCTCAGCACTCTCTTCAAAGCTTGCAGATTTTCAGTTCTTCATC	1572
			Db		
			2516	AGCAATGATGATGCTTTCGTTTAAAGGAAGAAATTTATTCATATTAAAGCAGATAGGAAG	2575
			Qy		
			1573	AATAATGATGATGCTTTCAGTTTAAAGGAAGAAATTTATTCATATTAAAGCAGATAGCAG	1632
			Db		
			2576	TGGAGGTTCAAGCAAGGATTTTCAGGTGTTTAAATGAAAAGAAACAGATATATGCTATAAA	2635
			Qy		
			1633	TGGAGGTTCCAGTAAGGTGTTTCAGGTGTTTGAATGAGAAAAAAGCAGATATACCGCTATCAA	1692
			Db		
			2636	ATATGTGAACCTTAGAAGACAGCAATACCAACTCTTGTAGTTACCGGAACGAATAGC	2695
			Qy		

```

1693 ATATGTGAACCTAGAGCGCGATAGCCAACTATTGAGAGCTACCGACGAGATAGC 1752
2696 TTATTTGAATAAATACAAACACAGATGATAGATCATCCGACTTTTATGATTATGAAT 2755
1753 GTTTTGAACAAACTACAGCAACACAGATGATAGATCATCCGCTCTATGATTATGAAT 1812
2756 CACGACACAGTACATCTACATGTAATGGAGTGTGGAATATTTGATCTTAAATAGTGGCT 2815
1813 CACGACAGTACATCTACATGTAATGGAGTGTGGAATATTTGATCTTAAATAGTGGCT 1872
2816 TAAAGAGAAAAATCCATTTGATCCATGGGAACGCAAGATTTACTGGAAAAATATGTTAGA 2875
1873 TAAAGAGAAAAATCCATTTGATCCATGGGAACGCAAGATTTACTGGAAAAATATGTTAGA 1932
2876 GGCAGTTACACAAATCCATTTGATCCATGGGAACGCAAGATTTACTGGAAAAATATGTTAGA 2935
1933 GGCAGTTACACAAATCCATTTGATCCATGGGAACGCAAGATTTACTGGAAAAATATGTTAGA 1992
2936 TCTGATAGTTGATGGAATGCTAAAGCTAAATGATTTTGGGATTCGCAACCAATGCAACC 2995
1993 TGTGATAGTTGATGGAATGCTAAAGCTAAATGATTTTGGGATTCGCAACCAATGCAACC 2052
2996 AGATACAAAGTGTGTTAAAGATTTCTCAGGTTGGCACAGTTAAATPATATGCCACCAGA 3055
2053 AGACACAAAGTGTGTTAAAGATTTCTCAGGTTGGCACAGTTAAATPATATGCCACCAGA 2112
3056 AGCAATCAAGATATGTTCTTCTCCAGAGAGATGGGAATCTCAAGTCAAGATAGACCC 3115
2113 AGCAATCAAGATATGTTCTTCTCCAGAGAGATGGGAATCTCAAGTCAAGATAGACCC 2172
3116 CAAAGTATGTTTGGTCTTAGATGATTTTGTACTATATGACTTACGGGAAAAACAC 3175
2173 CAGAAGTATGTTTGGTCTTAGATGATTTTGTACTATATGACTTACGGGAGGAGCC 2232
3176 ATTTTCAGCAGATATTAATCAGATTTCTAAATACATGCGCATTAATGATCTTAATCATGA 3235
2233 ATTTTCAGCAGATATTAATCAGATTTCTAAATACATGCGCATTAATGATCTTAATCATGA 2292
3236 AATTGAATTTCCCGATATTCAGAGAAAGATCTTCAAGATGCTTAAAGTGTGTTTAAA 3295
2293 GATTGAATTTCCCGATATTCAGAGAAAGATCTTCAAGATGCTTAAAGTGTGTTTAAA 2352
3296 AAGGACCCAAACAGAGGATATCCATTCCTGAGTCTCGGCTCATCATATGTTCAAT 3355
2353 GAGGAACCTTAAAGAGAGGATATCTATCCCTGAGTCTTCAACATCCGTTATGTTCAAT 2412
3356 TCAAACTCATCCAGTTAAACCAATGGCCAGGGAACCACTGAGAGAAATGAATATGTTCT 3415
2413 TCAGCCCATCCAGGAGCCAAATGGCTAGGAGGACCATGATGAATGAATATGTTCT 2472
3416 GGGCCAACTTGTGGTCTGAATTTCTCTAACTCCATTTTGAAGCTGCTTAAACCTTTATA 3475
2473 GGGTCAACTTGTGGTCTGAATTTCTCTAACTCCATTTTGAAGCTGCTTAAACCTTTATA 2532
3476 TGAACATATAGTGTGGTCAAGTATTAATTTCTCATCTCCAGACATTTTGAAGAAAA 3535
2533 TGAACGTTAATTTGGTGAAGGTCAAGATTTCTTGTATCCAGACATTTTGAAGAAAA 2592
3536 AAGGGAAAAAATGATTTTGCAGTTATTCGTATGTGCAGATAGGAGGTATATAATATAT 3595
2593 GAGAGAAAGAAAGTATGATGACACCTACGTACAAACCAAGA-----ACACTAGATGTT 2646
3596 GGAAGTTATCTTGAATCCCTGTGGAATCTA-CATTTGAAGACACATCATCTCTGA 3654
2647 CCTCTGCCATCTCTTGAATCTCTGAGGAATCTTACAGTTGGAAACCAACCTCAGCTGA 2706
3655 AGTGT-----ATCAGCAAAAAAATTCAGTGAAGATTTCTTTTAAAGAAAACT 3703
2707 TTTTATCAGTTAAAAAACAACAACAAACCTTCAGTAGATTTATCCCTCAAGAGAGCT 2766
3704 GTAAAAATAGCACCACTTATGGCAGCTGATATATTTGATAGATGTTTCT-CTGTTTAA 3762

```

```

Db 2767 GTAAAGTT---AACCACTCATAGCACTGTGTATATAATAATTATAGATGTGCTTTCTT 2823
QY 3763 TCTCTTGTGTAATCTACTTGACATCATTTTACTCTTGAATAGTGGGTG 3812
Db 2824 TTATGCTTTTCGTAAATCTGCTAATGTTTACGTTTAGACAGTGAATG 2873

RESULT 11
HSPYT HSPYT 946 bp mRNA linear PRI 30-JUN-1993
LOCUS H.sapiens mRNA for phosphotyrosine picked threonine kinase (PYT).
DEFINITION X70500 S53714
ACCESSION X70500.1 GI:312815
VERSION phosphotyrosine picked threonine kinase.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 946)
AUTHORS Lindberg,R.A.; Fischer,W.H. and Hunter,T.
TITLE Characterization of a human protein threonine kinase isolated by
screening an expression library with antibodies to phosphotyrosine
JOURNAL Oncogene 8 (2), 351-359 (1993)
MEDLINE 93149596
PUBMED 7678926
FEATURES
Location/Qualifiers
1..946
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
31..>946
/codon_start=1
/product="phosphotyrosine picked threonine kinase (PYT)"
/protein_id="CAA49912.1"
/db_xref="GOA:Q15272"
/translat="MCMRQLQVLASSANECISVKGRISILKQIGSGGSKVFOVL
NEKKQIVAIKVLNEADNOLDSYRNEIAYLNKLOHSDKIIRLYDYETTDQYIYVW
MECNINILNSWLKKSIDPWERKSYKKNLEAVHTIHOHGIHVSOLKPNPFLIVGM
LKLIDFGIANQMOPDTTSYVKDSQVGTVNYMPPEAIKDMSSRENGSKSKISPKSDV
WSUCILYITYGRTTPTQQLINQISLUHAIDENHEIEFPDIPEKDLQDVVKCKLRD
PKRISIPELLAHPYVQIQTHPVNQAGI"
ORIGIN
Query Match 23.2%; Score 895.4; DB 9; Length 946;
Best Local Similarity 99.9%; Pred. No. 1.3e-177;
Matches 896; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2487 CAAATTTACAGGTTTGTAGCATCTTCTCAGCAAAATGAATGCAATTCGTTAAAGGAAGA 2546
Db 43 CAAATTTACAGGTTTGTAGCATCTTCTCAGCAAAATGAATGCAATTCGTTAAAGGAAGA 102
QY 2547 ATTTATTTCCATATTAAGCAGATAGGAGTGGAGTTTCAGCAAGGTATTTTCAGTGTGA 2606
Db 103 ATTTATTTCCATATTAAGCAGATAGGAGTGGAGTTTCAGCAAGGTATTTTCAGTGTGA 162
QY 2607 AATGAAAGAAACAGATATATGCTATAAAAAATATGTGAACCTTAGAAGACAGATAACCAA 2666
Db 163 AATGAAAGAAACAGATATATGCTATAAAAAATATGTGAACCTTAGAAGACAGATAACCAA 2726
QY 2667 ACTCTTCATAGTTACCGAACGAAATAGCTTATTTGAACTTAACTACAAACAGTGTAT 2726
Db 223 ACTCTTCATAGTTACCGAACGAAATAGCTTATTTGAACTTAACTACAAACAGTGTAT 282
QY 2727 AAGATCATCCGACTTTATGATTAATGAATCAGCGACAGTACATCTACATGTTAATGGAG 2786
Db 283 AAGATCATCCGACTTTATGATTAATGAATCAGCGACAGTACATCTACATGTTAATGGAG 342
QY 2787 TGTGGAATATTTGATCTTAATAGTTGGCTTAAAGAAAAAATCCATTCATCCATGGGAA 2846
Db 343 TGTGGAATATTTGATCTTAATAGTTGGCTTAAAGAAAAAATCCATTCATCCATGGGAA 402

```

QY 2847 CGAAGAGTTACTGGAAAAATATGTTAGAGCGAGTTTCACAAATCCATCAACATGCGATT 2906
Db 403 CGAAGAGTTACTGGAAAAATATGTTAGAGCGAGTTTCACAAATCCATCAACATGCGATT 462
QY 2907 GTTCACAGTGAATCTTAACACGAGTAACTTCTGATAGTGTGATCGAATGCTAAAGCTAATT 2966
Db 463 GTTCACAGTGAATCTTAACACGAGTAACTTCTGATAGTGTGATCGAATGCTAAAGCTAATT 522
QY 2967 GATTTTGGGATTGCAAAACCAATGCAACCAAGATACACAAAGTGTGTTAAAGATTCTCAG 3026
Db 523 GATTTTGGGATTGCAAAACCAATGCAACCAAGATACACAAAGTGTGTTAAAGATTCTCAG 582
QY 3027 GTTGGCAGATTAATATATGTCACACAGAGCAATGCAAGATATGCTTCTCCACAGAG 3086
Db 583 GTTGGCAGATTAATATATGTCACACAGAGCAATGCAAGATATGCTTCTCCACAGAG 642
QY 3087 AATGGGAAATCTAAGTCAAGATGAAGCCCAAAAGTGAATGTTGGTCTTAGGATGTAAT 3146
Db 643 AATGGGAAATCTAAGTCAAGATGAAGCCCAAAAGTGAATGTTGGTCTTAGGATGTAAT 702
QY 3147 TTGACTATATGATCTTACGGGAAACACCAATTTCCAGCAGATTAATTAATCGATTTCTAAA 3206
Db 703 TTGACTATATGATCTTACGGGAAACACCAATTTCCAGCAGATTAATTAATCGATTTCTAAA 762
QY 3207 TTACATGCCATATGATCTTATCATGAATGAATTTCCCGATATCCAGAGAAAGAT 3266
Db 763 TTACATGCCATATGATCTTATCATGAATGAATTTCCCGATATCCAGAGAAAGAT 822
QY 3267 CTTCAAGATGTTTAAAGTGTGTTTAAAGGAGCCCAAAACAGAGATATCCATTCCT 3326
Db 823 CTTCAAGATGTTGTTAAAGTGTGTTTAAAGGAGCCCAAAACAGAGATATCCATTCCT 882
QY 3327 GAGCTCTGGTCTATCCATATGTTCAATTCAACTCATCCAGTTACCAAAATGGCC 3383
Db 883 GAGCTCTGGTCTATCCATATGTTCAATTCAACTCATCCAGTTACCAAAATGGCC 939

RESULT 12

AC112218 144319 bp DNA linear PRI 24-MAY-2002
LOCUS
DEFINITION Homo sapiens chromosome 3 clone CTD-2036B16, complete sequence.
AC112218
VERSION
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Kaul, R.K.; Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Phelps, K.A., Buckley, D., Raymond, C. and
Haugen, E.D.
Direct Submission
Unpublished
2 (bases 1 to 144319)
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
Direct Submission
Submitted (20-FEB-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 144319)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Phelps, K.A., Buckley, D., Raymond, C. and
Haugen, E.D.
Direct Submission
Submitted (24-MAY-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On May 24, 2002 this sequence version replaced gi:18767517.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: <http://www.genome.washington.edu>
Contact: uwachts@u.washington.edu
----- Project Information

Center project name: chr-3
Center clone name: CTD-2036B16 (bc0724)
----- Summary Statistics
Sequencing vector: plasmid; 100% of reads
Chemistry: Dye-terminator ET; 57% of reads
Chemistry: Dye-terminator Big Dye; 43% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 144239 bases at least Q40
Consensus quality: 144315 bases at least Q30
Consensus quality: 144319 bases at least Q20
Insert size: 144319; sum-of-contigs
Quality coverage: 10.3x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5': CTD-2200P7 (UWGC:bc0739)
3': Mapping in progress

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector. In order to accurately represent the entire circular BAC, small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

EcoRI				BglII				HindIII			
SeqDerMap	FngerPrint	SeqDerMap	FngerPrint	SeqDerMap	FngerPrint	SeqDerMap	FngerPrint	SeqDerMap	FngerPrint	SeqDerMap	FngerPrint
3200	3234	4451	4381	7510	7532						
6633	6881	2067	2050	19977	19743						
75	<800	3747	3776	3009	3037						
13226	12873	93	<800	8698	8821						
12985	12873	8674	8581	378	<800						
2918	2905	1940	2050	4338	4345						
17848	18128	9569	9330	8944	8821						
2843	2905	755	746	1891	1899						
10469	10380	3577	3561	2502	2484						
1195	1188	1018	1038	408	<800						
1938	1918	2555	2620	1361	1345						

Db 30583 ATTTGGTTTAAATTAATCTAAATATCTAAACATTTTGGACATTTGTTGATGTG 30642
Cy 900 AATGTAAGACTGTACAGACTTCTTAGAAAACAGTTTGG 937
Db 30643 AATGTAAGACTGTACAGACTTCTTAGAAAACAGTTTGG 30680

RESULT 13

AC097015 LOCUS AC097015 178596 bp DNA linear PRI 31-OCT-2002

DEFINITION Homo sapiens chromosome 3 clone RP11-82B23, complete sequence.

ACCESSION AC097015 AC023097

VERSION AC097015.2 GI:24431663

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE 1 (bases 1 to 178596) Sukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

2 (bases 1 to 178596) Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

3 (bases 1 to 178596) Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,

4 (bases 1 to 178596) Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and

5 (bases 1 to 178596) Haugen, E.D.

6 (bases 1 to 178596) Direct Submission

7 (bases 1 to 178596) Unpublished

8 (bases 1 to 178596) Kaul, R.K., Olson, M.V., Raymond, C., Clendenning, J. and Haugen, E.D.

9 (bases 1 to 178596) Direct Submission

10 (bases 1 to 178596) Submitted (05-OCT-2001) Genome Center, University of Washington,

11 (bases 1 to 178596) Box 352145, Seattle, WA 98195, USA

12 (bases 1 to 178596) Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,

13 (bases 1 to 178596) Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and

14 (bases 1 to 178596) Haugen, E.D.

15 (bases 1 to 178596) Direct Submission

16 (bases 1 to 178596) Submitted (31-OCT-2002) Genome Center, University of Washington,

17 (bases 1 to 178596) Box 352145, Seattle, WA 98195, USA

18 (bases 1 to 178596) On Oct 31, 2002 this sequence version replaced gi:15963672.

19 (bases 1 to 178596) Center: University of Washington Genome Center

20 (bases 1 to 178596) Center Code: UWGC

21 (bases 1 to 178596) Web site: <http://www.genome.washington.edu>

22 (bases 1 to 178596) Contact: uwgchg@u.washington.edu

23 (bases 1 to 178596) Drafting Center: BCM

24 (bases 1 to 178596) Project Information

25 (bases 1 to 178596) Center project name: chr-3

26 (bases 1 to 178596) Center clone name: RP11-82B23 (bc0204)

27 (bases 1 to 178596) Summary Statistics

28 (bases 1 to 178596) Sequencing vector: unknown; 54% of reads

29 (bases 1 to 178596) Chemistry: Dye-terminator Et; 87% of reads

30 (bases 1 to 178596) Assembly program: Phrap; version 0.990319

31 (bases 1 to 178596) Consensus quality: 178532 bases at least Q40

32 (bases 1 to 178596) Consensus quality: 178585 bases at least Q30

33 (bases 1 to 178596) Consensus quality: 178590 bases at least Q20

34 (bases 1 to 178596) Insert size: 178596; sum-of-contigs

35 (bases 1 to 178596) Quality coverage: 9.9x in Q20 bases; sum-of-contigs

36 (bases 1 to 178596) Overlapping Sequences:

37 (bases 1 to 178596) 5': RP11-5017 (UWGC:bc0107) AC096887, 10069-bp overlap

38 (bases 1 to 178596) 3': CTD-2036B16 (UWGC:bc0724) AC112218, 48985-bp overlap

39 (bases 1 to 178596) Sequence Quality Assessment:

40 (bases 1 to 178596) This entry has been annotated with sequence quality

41 (bases 1 to 178596) estimates computed by the Phrap assembly program.

42 (bases 1 to 178596) All manually edited bases have been reduced to quality zero.

43 (bases 1 to 178596) Quality levels above 40 are expected to have less than

44 (bases 1 to 178596) 1 error in 10,000 bp.

45 (bases 1 to 178596) Base-by-base quality values are not generally visible from the

46 (bases 1 to 178596) GenBank flat file format but are available as part

47 (bases 1 to 178596) of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest
fingerprinting. Comparison of the experimentally derived digest
fragments with sequence-predicted fragments is given below.
The electronically-digested sequence consists of both insert and
vector, in order to accurately represent the entire circular BAC.
Small fragments below a variable cutoff (approximately 400-800 bp)
are not resolved in the fingerprint and hence do not appear
in the table. There are no significant remaining discrepancies
between the experimental and predicted values. Uniquely ordered
fragments are separated by dashed lines.

ECORI				BglII			
SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
8696	8814	637	<800	8289	8317		
6	<800	6382	6511	2067	2009		
464	<800	512	<800	6977	6977		
9599	9606	449	<800	463	<800		
2859	2917	4961	4977	3522	3548		
17808	17859	165	<800	594	<800		
11932	11872	721	722	2615	2635		
10369	10364	4974	4977	1264	1261		
420	<800	689	<800	1058	1047		
3543	3585	2784	2825	891	890		
2337	2324	3624	3695	8871	8934		
2542	2579	846	866	5180	5210		
1961	1961	10774	10849	1898	2009		
2296	2324	164	<800	2223	2254		
414	<800	15284	15120	1428	1415		
50	<800	2753	2825	12404	12299		
5502	5501	3999	4079	14112	14052		
2914	2917	9649	9718	797	810		
1135	1115	4026	4079	6867	6977		
6215	6235	1984	2005	6362	6360		
1683	1687	446	<800	1966	2009		
24354	24444	3666	3695	1455	1415		
14824	14615	318	<800	2666	2635		
872	874	19350	19040	707	<800		

TITLE Mps1 is a kinetochore-associated kinase essential for the vertebrate mitotic checkpoint
 JOURNAL Cell 106 (1), 83-93 (2001)
 MEDLINE 21354313
 PUBMED 11461704
 REFERENCE 2 (bases 1 to 2649)
 AUTHORS Magnaghi-Jaulin,L., Vigneron,S., Lorca,T. and Labbe,J.C.
 TITLE X-Mps1/TTK is a novel Xenopus protein required for the spindle checkpoint
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 2649)
 AUTHORS Magnaghi-Jaulin,L., Vigneron,S., Lorca,T. and Labbe,J.C.
 TITLE Direct Submission
 JOURNAL Submitted (23-MAR-2000) CRBM, CNRS, 1919 route de Mende, Montpellier cedex 34293, France
 FEATURES Location/Qualifiers
 source 1..2649
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 CDS 1..2649
 /codon_start=1
 /product="Mps1/TTK"
 /protein_id="AAK27843.1"
 /db_xref="GI:13491131"
 /translation="MODELSERKLIASILDRVKFKYKYGDDNWTDELTFKSSA
 DTTEHSGIFHLVTTTPTEWLNCLKLENTGLPQIDPOLNKLIDNYSQAVGALPAE
 KHGHNESYAKILVHFAELKAHDLDEAREQFQRLNCKKFAFVHTFAQFELSEGN
 KKQKILRLQRCBEGCAYCEMLNALKNVQFKPOLISDEDKENLAVSSHIHQGTASF
 ONLAGNPRQMKTESPEEYKTRFSYGEKLSPEDEIDGRKPLLNMAKTCPLGRV
 PVQAPSPDRTKRSKDGSSSVK3PFSMRVPVLATISNPKYLGDDIDCHSEDIVP
 SSNLPHEEONANEDSLDKTPSSVILSVTESNLTVKNEDDLILGSTSTAVLPQENSK
 PSESAPLLENINARSLDITKPNNEAEKWKVETPRQIFQPEMIFPAKRRKSTEPS
 ANFGRSRVSPAPSLKCDKDFVFCGTVFNKSDQYDNCFTPVKTNLGLAQMSTPY
 NNRSDYQPTQPGQDCPFIAPFASDCIVVGRAYALVKQIGTSSKVFQWMD
 KKHLYALVYNLEADQQTIESYQNEISHLNKLQCHDKIIRLFDEIYDQHIYMWY
 CGNIDLNLWRKKTNIPKRSYKWNMLEAVHTIHCNGIVHSDLPANFLIVDGLK
 LIDFGANQIPQVTSIKVDSQVGTINYPPEISRTTSYAENGKPSRSKPGDWVS
 LGILCYMTTGTTPFOHITNQIAKLSHILDPGYEIEFPNIPKDLQDVLKRLKLVNPK
 ERISIAELLVHPVYIQPHTPQDQVQKQTEEMKRIILGQILGLNSFNSISRAAKNLY
 DQPNSGRSLSLTLGTVTSQNTRTTK"

ORIGIN

Query Match 20.5%; Score 793.8; DB 5; Length 2649;

Best Local Similarity 60.2%; Pred. No. 2.6e-156;

Matches 1575; Conservative 0; Mismatches 947; Indels 93; Gaps 12;

978 ATGGAATCCGAGGATTTAAGTGGCAGAGAATTGCAATTCATTCATTAATGAACAAGTG 1037

1 ATGGATGATGAAGATATAAGTGAACGGAACCTGAAATTCGTTCTATCTTGGATAGATT 60

1038 AGACACATTAATAATAGTTTAAATGAAGACCTT---ACTGATGACTAATTCATTGAAT 1094

61 AAGAGTTTAAAGACCAAGTATGGAACAGATGATAATTTGGACTGATGAATTAACCTTTCA 120

1095 AAAATTTCTGCTGATACACAGATAACTCGGGAACCTGTAACCAAAATTATGATGATGGCA 1154

121 AAGTCTTCTGCTGACATACAGACATCTGGCATTTTACCCA---CTTGTGTGACCACA 177

1155 AACACCCAGAGACTGGTGTAGTTTGTGTCATAACTAGAGAAAACAGTGTTCGCTTA 1214

178 AAGACTCTGAGGAATGGCTAAATTTGCTTCTTAAATTTAGAAAATACAGGCCTGCCTCAG 237

1215 AGTCATGC---TCTTTTAAATAAATTCATTGGTGTGCTGTACAGTCAAGCAATTGAAGCGCTT 1271

238 ATTGACCCACACTTCTCAATAGCTCATTTGATATTAATACAGTGTGAGGACCTTG 297

1272 CCCCAGATAAATATGCCAAAATGAGAGTTTGTCTAGAAATTCAGTGAGATTGTGTAA 1331

298 CCCGACAGAAAACACAGTCAATGAGAGCTATGCTAAATCTCTGTGCACTTTGCGGAA 357

1332 TTAAGACTATTCAAGAGCCAGATGATGCACGTGACTACTTTCAATGTGGCCAGCAAC 1391

Db 358 TTGAAGCAATCCATGACCTAGATGAAGCAGGGCAGATTTCAGTTTGCAGACTAAAC 417

Qy 1392 TGCAGAGAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAACCTGTCAACAGTAAAT 1451

Db 418 TGCAGAAAGTTTGTCTTTTGTGCATACAGCTTTTGCCTAATTTGAGCTGTGAGAGGAAAT 477

Qy 1452 GTCAAAAAAAGTAAACAACTTCTTCAAAAAGCTGTAGAACCTGTGAGCAGTACCTACTAGAA 1511

Db 478 TTTAAAAAGTGTAGCAGATTCTTGAAGAGGTCTGGAATGTGGCGCTGTTCCTCCGCGAA 537

Qy 1512 ATGCTGGAAATTCCTCTCGGGAATTTAAACCTCCAAAAAAGCAGCTGTCTTTCAGAGGAG 1571

Db 538 ATGTTGAACTTTGGCTTTTGAAGAAATGTACAGTTTAAAGCCTCAGCTGATATCTGACGAA 597

Qy 1572 GAAAAAGAAAGATTTATCAGCATC-TACGGTATTAACTGCCCAAGAAATCATTTTCGGGTTTC 1630

Db 598 GATAAAGAAATTTAGCAGTGTCTCATCAAGTCAATCAATCAAGGACGGCTAGTTTTCAG 657

Qy 1631 ACTTGGGCAATTTACAGATAGGAACACAGTTGTGATT-----CCAGAGGACAGACTACT 1685

Db 658 AATTTGGCACTGGGAAATCCACAGAAATGAAGATCGAAAGCCCAAGAAATATTTCTGTA 717

Qy 1686 AAAGCCAGGTTTATATGAGAGAAACATGCCACCAAGATGCAGAAATAGGTTTACCGG 1745

Db 718 AAAACAAGATTTTATATGAGAAAGTTAAGCTCCCAGAAAGACTTTGAAGATATAGTA 777

Qy 1746 AATTCATTGAGCAAACTTAAACAACTAAACAGTCATGCCATTTGGGAAGAGTCCAGTT 1805

Db 778 AGGAAACCGTTACT---TAACATGTCTGCTAAGAACCTCTCTCTTTGGGAGAGTGCCTGTA 834

Qy 1806 AACCTTCTAAATAGCCCAAGT---TGTGATGTGAAGACAGATGATTTCAGTTTGTACTTGT 1862

Db 835 CAAACAGCAATCTTCAGATACAGAAACAGGAAGAGTGTGCTCTGGCAGATTCAGT 894

Qy 1863 TTTATGAAAAGCAAAACCTCTAGATCAAGATCCGAGATTTGGTTGTGCCCTGGATCTAAA 1922

Db 895 GTTGTGAAGAGACCAATTCACTCCATCCATGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCT 954

Qy 1923 CCAAGTGAATGATTCCTGTGAAATTAAGAAATTAAGATCTGTTTCAAAATAGTCAATTC 1982

Db 955 AAATATTTGGGAGATGATCTGCAATTCAGAAAGACATTAAGAGTACCAGAGCAACATC 1014

Qy 1983 AAGGAACCTCTGGTGTGATGAAAGAGTTCTGAACTTATTATTACTGATTCAATAACC 2042

Db 1015 CTACCATGATGAAGACAGCTTAATGAAGACAGCTTAGATATGAAGACACCGTCTCTCT 1074

Qy 2043 CTGAGAAATAAAACGAATCAAGTCTTCTGCTAAATTAGAAGAA----- 2087

Db 1075 ATTCTTAGTGTACAGAAATCAAACTGACTGTCTCAAAAGAAATGAAGATTTAATTTGGGA 1134

Qy 2088 -----ACTAAAGAGTATCAAGAACCAAGAGGTTCCAGAGAGT 2123

Db 1135 AGCAGCTCTACAGTGTCTCAATCTCAAGAAACAGTAAACCATCTGAATCAGCACCA 1194

Qy 2124 AACCAAGAACAGTGGCAAGCTAAAGAAAGTCAAGTGTATTAAACGAAATCTCTGCTGCA 2183

Db 1195 CCACCTTTTACCCACCAATTAATGAAGGCTCTGACACGACCAAGCCAAACAATATTGAA 1254

Qy 2184 TCTTCAATCACTGCGAGATTCGCGAGTTAGCCGAAAGTT----- 2225

Db 1255 CAGAAATGGAAGTGAAGGTTCCAGAGACACCCAGGCAAAATTTCCAGCCAGAGATGATA 1314

Qy 2226 AATACAGAGCAGAAACATACCACTTTTGAAGCAACCTGTCTTTTTCAGTTTCAAAACA---G 2282

Db 1315 TTTCCAGCGGTAAAAGAAAGAGTACTGAGCTAGTGTCTGCTGTTTCCGACGAGTG 1374

Qy 2283 TCACACCAATATCAACTTAAATGTTTACCCAAATCTATTGTTGAAGACACCAAGC 2342

Db 1375 TCTCGCCGACGACCTCTCTCTTAGTAAATGTGACCCCTGTCTTTGTTGGAACACCACTC 1434

Qy 2343 AGCAATCTCTGGATGATTACATGAGCTGTTTGAAGTCCAGTTTGAAGAAATGACCTTT 2402

Db 1435 AATAATCACAAGAGATTTACATGACTGCTTTAGACTCTCAGTTTGAAGAAACGACCTA 1494

2403 CCACCTGCTTGTGAGTGTCAACACCTTATGGCAACCTGCGCTTTTCAGGAGCAACAG 2462
 Db GGACCATTTGGCACAATAATGTCACACCTTACACCAACCGTTTCAGATTATCAGCAACCCACAG 1554
 2463 CATCAAACTATGTCACCTCCACCTTCAAAATTTACAGGTTTATAGCATCTTCTCAGCAAT 2522
 Db ACCCAGTTGGA-----CAACCGGACTGTTTCCAACTCTGCACTTTTGCATCAAGT 1608
 2523 GAATGCACTTTCGTTTAAAGGAAGAATTTATTCATATTAAGCAGATAGGAAGTGGAGT 2582
 Db GACTGATTGTTTAAAGGACGTGCATATGCACTTTTAAACCAAAATGGAACAGGTGA 1668
 2583 TCAAGCAAGTATTTCAAGTGTGTTAAATGAAGAAACAGATATATCTATATAAATATGTG 2642
 Db TCTAGTAAGTGTGTTTCAAGTAATGGAATGAAGAAACATTTATATGCCATAAGTATGTG 1728
 2643 AACTTGAAGAAGCAGATACCAAACTCTGTATAGTTTACCGAAGCAAAATAGCTTATTTG 2702
 Db AATCTGARGAGCTGACACGAGCAGACCATAGAAAGCTATCAGATGAATTTCTCACITG 1788
 2703 AATTAACATCAACACACGATGATAGATCATCCGACTTTTATGATTAATGAATCAGGAC 2762
 Db AACAACCTTTCAGCAACACTGTGCAAGATCATTAGGCTCTTTGACTATGAGATTACTGAC 1848
 2763 CAGTACATCTACATGTTGGAAGTGTGGAATATTTGATCTTAATAGTTGGCTTAAAG 2822
 Db CAGCATATTTACATGTTGGAAGTGTGGAATATAGATCTCAATCTTGGCTTAGGAAG 1908
 2823 AAAAACTCCATTCATGCGGAACGACAGATTTACTGGAATAATATGTTAGAGGAGTT 2882
 Db AAAAACTATTAATCTTGGGAACGAAAGAGCTACTGGAATAATATGTTGGAAGCAGTC 1968
 2883 CACACATCCATCAACATGCTATGTTTACAGTATCTTAAACCCAGTAACTTTCTGATA 2942
 Db CACACATATCATCAGATGGATTTGACAGATGCTTGAAGCCAGCAACTTTCTAAT 2028
 2943 GTTATGGAATGCTAAAGCTAATTTGATTTTGGGATTCGAAACCAATGCAACAGATACA 3002
 Db GTTACGGGATGCTCAAGCTAATAGACTTTGGCATAGCAACCAATCAGCCAGATGTA 2088
 3003 ACAAGTGTGTTTAAAGATCTCAGTGTGGCAGATTTAATATATGCCACAGAGCAATC 3062
 Db ACCAGCATTTTAAAGATTCACAGTGTGGCAGATTAATATATGCCACAGCAATCTATC 2148
 3063 AAAGATATGCTTCTCCAGAGAGATGGGAATCTAAGTCAAGATAGCAAGCCCAAAAGT 3122
 Db AGAGATACCACTCATATGCGGAATGCGAATCCAGATCAAGATTAGCCCAAGAGT 2208
 3123 GATGTTGTCCTTAGGATGATTTTGTACTATATGACTTACGGGAAACACCAATTCAG 3182
 Db GATGTCGTGTCCTGGGTTGCATATTTGTACTGATGACATATGGGAAGACTCTTTTCAA 2268
 3183 CAGATAATTAATCAGATTTCTAATATACATGCTAATGATCTTAACTCATCAATGAAATGAA 3242
 Db CATATTACAATCAGATGCTAATCTCAATTTCTGATCTGATGTTTAAAGATTAGAG 2328
 3243 TTTCGGATATTCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAGAGGAC 3302
 Db TTCCCTAATATACAGAGAAAGACCTTCAAGATGTTTAAAGATGTTTAAAGATGTTTAA 2388
 3303 CCAAAACAGAGGATATCCATCTGAGCTCTGCTCATCATCTGATCTCAAT-----T 3356
 Db CTTAAGGAAGGATATCCATTTGCTGAGCTCTTGTGACCCCATATGTTTCAAGATCAGCT 2448
 3357 CAAACTCATTCAGTTAAACCAATGGCCAAAGGAAACCACTGAAGAAATGAATATGTTCTG 3416
 Db CATACACAGCCGATCAACAAAGTGAAGGAAAGAGACGACAGAGAAATGAAGGATATCTC 2508
 3417 GGCACACTTGTGCTGATCTCTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 3476
 Db GGACGCTTATTTGCTGATTTCTCAATTTCCATATCCGAGCTGCAAAAAAATTTGAT 2568

3477 GAACACTATAGTGTGGTGAAGTCAATATCTTC 3511
 Db GATCAGTTTACAGTGGCAGAGCCTGGATCTGTC 2603

RESULT 15
 AC022415/15
 LOCUS Homo sapiens chromosome 19 clone CTC-359D24, complete sequence.
 AC022415
 AC022415.5 GI:11968305
 VERSION HTG.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 180510)
 DOE Joint Genome Institute and Stanford Human Genome Center.
 REFERENCE
 AUTHORS Direct Submission
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 180510)
 REFERENCE
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (03-FEB-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 3 (bases 1 to 180510)
 DOE Joint Genome Institute and Stanford Human Genome Center.
 REFERENCE
 AUTHORS Direct Submission
 TITLE Submitted (22-DEC-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 JOURNAL On Dec 22, 2000 this sequence version replaced gi:7704995.
 COMMENT Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov
 Finishing Completed at Stanford Human Genome Center
 www.shgc.stanford.edu
 Quality: Phrap Quality >=40 99.9% of Sequence;
 Estimated Total Number of Errors is 0.1.
 STS Content:
 SHGC-6150 G14148.
 Location/Qualifiers
 source 1..180510
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="19"
 /clone="CTC-359D24"

Query Match 10.8%; Score 418; DB 9; Length 180510;
 Best Local Similarity 81.1%; Pred. No. 1.8e-77;
 Matches 498; Conservative 0; Mismatches 115; Indels 1; Gaps 1;

9 TTTTATTTTATTTTTCAGATGAGTTTCACTCTTTGTCAGGCTGGAGTGAATGGCA 68
 Db TTTTATTTTATTTTTCAGAGCGAGTTTCACTCTTTGTCAGGCTGGAGTGAATGGTG 129490
 69 CAATCTCAGTTTACTGCAACTCCGCTCCGCGTTCAAGCGATTCTCTGCTCAGCCT 128
 Db CAATCTCAGTCACTCCGCAATCTCTGCTCCGCGTTCAAGTGAATCTCTGCTCAGCCT 129430
 129 CTCAAGTAGCTGGATTACAGGATGTGCCACCACTGGCTTAACCTTTCTTTCTA 188
 Db CCCAGTAGCTGGATTACAGGATGTGCCACCACTGGCTTAACCTTTCTTTCTA 129370
 189 TTTAGTAGAGATGGGTTTCCATGTTGTCAGGCTGGTCTTGAATCTGACCTCAGG 248
 Db TTTAATAGACAGAGGTTTCTCCATGTTGTCAGGCTGGTCTTGAATCTGACCTCAGG 129310
 249 TGATCCACTTTGCTGCTCCCAAGAGTGTAGATTTACAGCGCTGAATGTCGCTGCG 308
 Db TGATCCGCCACCTTGGCTTCCCAAGAGTGTGGATTTAGGATGAGGCACTGCGCTG 129250
 309 TGATCTCTTTTGTGTTGATTTTGAACAGGGTCTCCCTTTGTCGCCCCAGGCTGGA 368

[illegible]

Search completed: September 30, 2004, 06:35:00
Job time : 14333 secs

This Page Blank (uspto)

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	3866	100.0	3866	9	US-09-967-760A-148		Sequence 148, App
2	3866	100.0	3866	13	US-10-342-887-899		Sequence 899, App
3	3866	100.0	3866	13	US-10-172-118-899		Sequence 899, App
4	3866	100.0	3866	15	US-10-081-119-13		Sequence 13, App
5	3866	100.0	3866	15	US-10-351-358-29		Sequence 29, App
6	3866	100.0	3866	15	US-10-116-712-663		Sequence 663, App
7	3866	100.0	3866	16	US-10-353-690-1		Sequence 1, Appl
8	3866	100.0	3866	16	US-10-366-288-15		Sequence 15, App
9	3866	100.0	3866	16	US-10-295-027-803		Sequence 803, App
10	3866	100.0	3866	16	US-10-298-027-832		Sequence 832, App
11	3866	100.0	3866	16	US-10-299-027-864		Sequence 864, App
12	2569.2	66.5	2574	16	US-10-173-999-33		Sequence 491, App
13	2569.2	66.5	2574	16	US-10-173-999-33		Sequence 33, App
14	402	10.4	2470	13	US-10-021-632-110237		Sequence 110237, App

QY 121 CTCAGCCTCTCAAGTAGTGGGATTACAGGCATGTGCCACCCCTGGCTAACTAATTT 180
Db 121 CTCAGCCTCTCAAGTAGTGGGATTACAGGCATGTGCCACCCCTGGCTAACTAATTT 180
QY 181 CTTTTCTATTAGTAGAGATGGGGTTTACCATGTTGGTCAGGCTGGTCTTTGAACCTCTG 240
Db 181 CTTTTCTATTAGTAGAGATGGGGTTTACCATGTTGGTCAGGCTGGTCTTTGAACCTCTG 240
QY 241 ACCTCAGGTGATCCACTTGGCTTGGCTCCAAAGTCTAGGATTAACGCCGTGAAACTG 300
Db 241 ACCTCAGGTGATCCACTTGGCTTGGCTCCAAAGTCTAGGATTAACGCCGTGAAACTG 300
QY 301 TGCCTGGCTGATCTTTTTTTTGTGTTGGATTTTGAACACAGGCTCTCCCTGGTGGCC 360
Db 301 TGCCTGGCTGATCTTTTTTTTGTGTTGGATTTTGAACACAGGCTCTCCCTGGTGGCC 360
QY 361 AGGCTGAGTGCAGTGGTGGATCTTGGCTCACTATAACCTCCAGCTCTGGTTCAAGT 420
Db 361 AGGCTGAGTGCAGTGGTGGATCTTGGCTCACTATAACCTCCAGCTCTGGTTCAAGT 420
QY 421 GATCCTCCCACTTTAGCCTCTCAGTAGCTGTGATTACAGCGGTGACACACACCCGG 480
Db 421 GATCCTCCCACTTTAGCCTCTCAGTAGCTGTGATTACAGCGGTGACACACACCCGG 480
QY 481 CTAATTTTTGTATTTTATTAGACACAGGTTTCCACATGTTGGCCAGGCTGTCTCAAA 540
Db 481 CTAATTTTTGTATTTTATTAGACACAGGTTTCCACATGTTGGCCAGGCTGTCTCTCAA 540
QY 541 CTCTGACCTCAAGGATCGCTGCTCCACTCCCAAGTCCCGAGATTACAGTGTG 600
Db 541 CTCTGACCTCAAGGATCGCTGCTCCACTCCCAAGTCCCGAGATTACAGTGTG 600
QY 601 AGTCACCATCCTGACCTTAATAATCTTAAGTCAATTTTTCTGGTCCATTTCTTCCCTAG 660
Db 601 AGTCACCATCCTGACCTTAATAATCTTAAGTCAATTTTTCTGGTCCATTTCTTCCCTAG 660
QY 661 GGTCTCCACACAAATCTGATTTAGGGGTGACAAATCTCTTAATCTTGATTCACAA 720
Db 661 GGTCTCCACACAAATCTGATTTAGGGGTGACAAATCTCTTAATCTTGATTCACAA 720
QY 721 AGGAAGTAGAAGTCAATCATGATTTAGAAAGGGGAAGTAGTAGGCCACTGCACACTCCT 780
Db 721 AGGAAGTAGAAGTCAATCATGATTTAGAAAGGGGAAGTAGTAGGCCACTGCACACTCCT 780
QY 781 GGATGATGATCCTAAATCCAGATACAGTAAATAATGGGGTATGGGAAGGTAGAAATCAAAA 840
Db 781 GGATGATGATCCTAAATCCAGATACAGTAAATAATGGGGTATGGGAAGGTAGAAATCAAAA 840
QY 841 TTTGGTTTAAATTAATTTATCTAAATATCTAAATAACATTTTGGATACATTTGATGTGA 900
Db 841 TTTGGTTTAAATTAATTTATCTAAATATCTAAATAACATTTTGGATACATTTGATGTGA 900
QY 901 ATGTAAGACTGTACAGACTTCTTAGAAAACAGTTTGGGTTCATCTTTTCCCTCCAG 960
Db 901 ATGTAAGACTGTACAGACTTCTTAGAAAACAGTTTGGGTTCATCTTTTCCCTCCAG 960
QY 961 TGCAGTTTCTGTAGAAATGGATCCGAGATTTAAGTGGCAGAGATTTGACAATGATT 1020
Db 961 TGCAGTTTCTGTAGAAATGGATCCGAGATTTAAGTGGCAGAGATTTGACAATGATT 1020
QY 1021 CCATAATGAACAAAGTGAGAGACATTTAAATAAAGTTTAAATAAGAACCTTACTGTATG 1080
Db 1021 CCATAATGAACAAAGTGAGAGACATTTAAATAAAGTTTAAATAAGAACCTTACTGTATG 1080
QY 1081 AACTAGCTTGATATAAATTTCTGTGATCTACAGATTAACCTCGGAACCTGTAAACAAA 1140
Db 1081 AACTAGCTTGATATAAATTTCTGTGATCTACAGATTAACCTCGGAACCTGTAAACAAA 1140
QY 1141 TTATGATGATGCAAAACAAACCCAGAGACTGGTTGAGTTTGTCTCAAACTAGAGAAA 1200
Db 1141 TTATGATGATGCAAAACAAACCCAGAGACTGGTTGAGTTTGTCTCAAACTAGAGAAA 1200
QY 1201 ACAGTGTTCGGCTAAGTAGTCTCTTTTAAATAAATTTGATTTGGTCGTTACAGTCAAGCAA 1260

Db 1201 ACAGTGTTCGGCTAAGTAGTCTCTTTTAAATAAATTTGATTTGGTCGTTACAGTCAAGCAA 1260
QY 1261 TTGAAGCGCTTCCCGCCAGATAAATATGGCCAAAATGAGAGTTTGTGTAGAAATTCAGTGA 1320
Db 1261 TTGAAGCGCTTCCCGCCAGATAAATATGGCCAAAATGAGAGTTTGTGTAGAAATTCAGTGA 1320
QY 1321 GATTGTGTAATTAAGACTATTCAAGAGCCAGATGATGACGTGACTACTTCAAAATGG 1380
Db 1321 GATTGTGTAATTAAGACTATTCAAGAGCCAGATGATGACGTGACTACTTCAAAATGG 1380
QY 1381 CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTCGAATGT 1440
Db 1381 CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTCGAATGT 1440
QY 1441 CACAAGGTAAATGTCAAAAAAGTAAACAACCTTTCTCAAAAAGCTGTAGAACGTGGAGCAG 1500
Db 1441 CACAAGGTAAATGTCAAAAAAGTAAACAACCTTTCTCAAAAAGCTGTAGAACGTGGAGCAG 1500
QY 1501 TACCACATAGAAATGCTGGAATTTGCCCTGCGGAATTTAAACCTCCAAAAGAGCAGCTGC 1560
Db 1501 TACCACATAGAAATGCTGGAATTTGCCCTGCGGAATTTAAACCTCCAAAAGAGCAGCTGC 1560
QY 1561 TTTCAGAGGAGAAAGAAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAGAAATCAT 1620
Db 1561 TTTCAGAGGAGAAAGAAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAGAAATCAT 1620
QY 1621 TTTCCGGTTCACTTTGGGCATTTTATATGGAGAGAAACATGCCACCAAGATGCAAGAAATAGGTT 1740
Db 1621 TTTCCGGTTCACTTTGGGCATTTTATATGGAGAGAAACATGCCACCAAGATGCAAGAAATAGGTT 1740
QY 1681 CTACTAAAGCCAGGTTTTTATATGGAGAGAAACATGCCACCAAGATGCAAGAAATAGGTT 1740
Db 1681 CTACTAAAGCCAGGTTTTTATATGGAGAGAAACATGCCACCAAGATGCAAGAAATAGGTT 1740
QY 1741 ACCGGATTCATTGAGACAAACTAAACAAACTAAACAGTCAATGCCATTTGGGAAGTGC 1800
Db 1741 ACCGGATTCATTGAGACAAACTAAACAAACTAAACAGTCAATGCCATTTGGGAAGTGC 1800
QY 1801 CAGTTAACTCTTAAATAGCCAGATTTGATGTGAAGACAGATGATTCACTTTGTACCTT 1860
Db 1801 CAGTTAACTCTTAAATAGCCAGATTTGATGTGAAGACAGATGATTCACTTTGTACCTT 1860
QY 1861 GTTTTATGAAAAGACAAACCTCTAGATCAGAAATGCCAGATTTGGTTGCTCGATCTA 1920
Db 1861 GTTTTATGAAAAGACAAACCTCTAGATCAGAAATGCCAGATTTGGTTGCTCGATCTA 1920
QY 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCAAT 1980
Db 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCAAT 1980
QY 1981 TCAAGAACCTCTGGTCTCAGATGAAAAGAGTTCTGAACCTTATTATTACTGATTCATAA 2040
Db 1981 TCAAGAACCTCTGGTCTCAGATGAAAAGAGTTCTGAACCTTATTATTACTGATTCATAA 2040
QY 2041 CCCTGAGAAATTAAGCGGAATCAAGTCTTCTAGCTAAATTTAGAGAACTTAAAGATGATC 2100
Db 2041 CCCTGAGAAATTAAGCGGAATCAAGTCTTCTAGCTAAATTTAGAGAACTTAAAGATGATC 2100
QY 2101 AAGAACCCAGAGGTTCCAGAGAGTAAACAGAAAACAGTGGCAAGCTAAGAGAAAGTCAAGT 2160
Db 2101 AAGAACCCAGAGGTTCCAGAGAGTAAACAGAAAACAGTGGCAAGCTAAGAGAAAGTCAAGT 2160
QY 2161 GTATTACAGAAATCTGTGCTCTTCAATCACTGGCAGATTCGGAGTTAGCCCGAA 2220
Db 2161 GTATTACAGAAATCTGTGCTCTTCAATCACTGGCAGATTCGGAGTTAGCCCGAA 2220
QY 2221 AAGTTAATACAGAGCAGAAACATACCACTTTTGCAGCAACCTGCTTTTCAAGTTCAAAAC 2280
Db 2221 AAGTTAATACAGAGCAGAAACATACCACTTTTGCAGCAACCTGCTTTTCAAGTTCAAAAC 2280
QY 2281 AGTCACCAACCAATATCAACATCTAAATGTTTGAACCCAAAATCTATTGTTAGACACCAA 2340

Db 2281 AGTCACACCAATATCAACATCTAAATGGTTTGACCCAAAATCTATTGTGAAGACACCAA 2340
Qy 2341 GCAGCAATACCTGGATGATTACATGAGCTGTTTTAGAACTCCAGTTGTAAAGAAATGACT 2400
Db 2341 GCAGCAATACCTGGATGATTACATGAGCTGTTTTAGAACTCCAGTTGTAAAGAAATGACT 2400
Qy 2401 TTCCACCTGCTGTGTCAGTGTCAACACCTTATGGCCAACTGCTGTTTTCCAGCAGCAAC 2460
Db 2401 TTCCACCTGCTGTGTCAGTGTCAACACCTTATGGCCAACTGCTGTTTTCCAGCAGCAAC 2460
Qy 2461 AGCATCAATATCTTGGCACTCCACTTCAAAAATTTACAGGTTTTAGCATCTTCTTCAGCAA 2520
Db 2461 AGCATCAATATCTTGGCACTCCACTTCAAAAATTTACAGGTTTTAGCATCTTCTTCAGCAA 2520
Qy 2521 ATGAATGCAATTCGGTTAAAGGAAGAAATTTATCCATATTAAGCAGATAGGAAGTGGAG 2580
Db 2521 ATGAATGCAATTCGGTTAAAGGAAGAAATTTATCCATATTAAGCAGATAGGAAGTGGAG 2580
Qy 2581 GTTCAAGCAAGGTATTTTCAGTGTAAATGAAGAAAGAACAGATATATGCTATATAATATG 2640
Db 2581 GTTCAAGCAAGGTATTTTCAGTGTAAATGAAGAAAGAACAGATATATGCTATATAATATG 2640
Qy 2641 TGAATCTAGAAAGAGAGATAACCAACTCTTGTAGTATCCGGAACGAATAGCTTAT 2700
Db 2641 TGAATCTAGAAAGAGAGATAACCAACTCTTGTAGTATCCGGAACGAATAGCTTAT 2700
Qy 2701 TGAATTAACATACACACAGTATAGATCATCCGACTTTATGATTAATGAATCAGCG 2760
Db 2701 TGAATTAACATACACACAGTATAGATCATCCGACTTTATGATTAATGAATCAGCG 2760
Qy 2761 ACCAGTACATCATAGTATGAGTGTGGAATATTTGATCTTAAATAGTGGCTTAAAA 2820
Db 2761 ACCAGTACATCATAGTATGAGTGTGGAATATTTGATCTTAAATAGTGGCTTAAAA 2820
Qy 2821 AGAAAAATCCATTTGATCCATGGGAACGACAGAGTTACTGGAAAAATATGTTAGAGCGAG 2880
Db 2821 AGAAAAATCCATTTGATCCATGGGAACGACAGAGTTACTGGAAAAATATGTTAGAGCGAG 2880
Qy 2881 TTACACAAATCCATCAATGGCATTTGTTTACAGTGTATTTAAACACAGCTTAACTTTCTGA 2940
Db 2881 TTACACAAATCCATCAATGGCATTTGTTTACAGTGTATTTAAACACAGCTTAACTTTCTGA 2940
Qy 2941 TAGTTGATGGAATCTTAAAGCTAATTTGATTTGGGATTTGCAACCAATGCAACCCAGATA 3000
Db 2941 TAGTTGATGGAATCTTAAAGCTAATTTGATTTGGGATTTGCAACCAATGCAACCCAGATA 3000
Qy 3001 CAACAAGTGTGTTAAAGATTTCTCAGTGTGACAGTAAATATATATCCACCAAGAGCAA 3060
Db 3001 CAACAAGTGTGTTAAAGATTTCTCAGTGTGACAGTAAATATATATCCACCAAGAGCAA 3060
Qy 3061 TCAAGAGATATGTTCTTCCAGAGAGAAATGGGAATCTAAGTCAAGATAAGCCCAAAA 3120
Db 3061 TCAAGAGATATGTTCTTCCAGAGAGAAATGGGAATCTAAGTCAAGATAAGCCCAAAA 3120
Qy 3121 GTGATGTTGCTGCTAGGATGTTTGTACTATATGACTTACGGGAACACCACTTTC 3180
Db 3121 GTGATGTTGCTGCTAGGATGTTTGTACTATATGACTTACGGGAACACCACTTTC 3180
Qy 3181 AGCAGATAATTAATCAGATTTCTAAATTACATGCCATAATGTATCCCTAATCATGAATG 3240
Db 3181 AGCAGATAATTAATCAGATTTCTAAATTACATGCCATAATGTATCCCTAATCATGAATG 3240
Qy 3241 AATTTCCGATATTTCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAGGG 3300
Db 3241 AATTTCCGATATTTCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAGGG 3300
Qy 3301 ACCCAAAACAGAGATATCCATCTCAGCTCTGCTGCTATCCATATGTTTCAAAATCCAA 3360
Db 3301 ACCCAAAACAGAGATATCCATCTCAGCTCTGCTGCTATCCATATGTTTCAAAATCCAA 3360
Qy 3361 CTCAATCCAGTTAAACAAATGCCAAGGGAACCACTGAAGAAATGAATATGTTCTGGGCC 3420
Db 3361 CTCAATCCAGTTAAACAAATGCCAAGGGAACCACTGAAGAAATGAATATGTTCTGGGCC 3420

RESULT 2

US-10-342-887-899

; Sequence 899, Application US/10342887

; Publication No. US20040058340A1

; GENERAL INFORMATION:

; APPLICANT: Dai, Hongyue

; APPLICANT: He, Yudong

; APPLICANT: Linsley, Peter S.

; APPLICANT: Mao, Mao

; APPLICANT: Roberts, Christopher J.

; APPLICANT: Van 't Veer, Laura Johanna

; APPLICANT: Van de Vijver, Marc J.

; APPLICANT: Bernards, Rene

; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

; FILE REFERENCE: 9301-188-999

; CURRENT APPLICATION NUMBER: US/10/342,887

; PRIOR FILING DATE: 2003-01-15

; PRIOR APPLICATION NUMBER: 60/298,918

; PRIOR FILING DATE: 2001-06-18

; PRIOR APPLICATION NUMBER: 60/380,710

; PRIOR FILING DATE: 2002-05-14

; PRIOR APPLICATION NUMBER: 10/172,118

; PRIOR FILING DATE: 2002-06-14

; NUMBER OF SEQ ID NOS: 2699

; SEQ ID NO 899

; LENGTH: 3866

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-342-887-899

Query Match 100.0%; Score 3866; DB 13; Length 3866;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3866; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAATTCCTTTTTTTTTTTTTTTTTTTTGGATGGAGTTTCACTCTGTTGGCCAGCTGGAGTG 60

Db 1 GGAATTCCTTTTTTTTTTTTTTTTTTTTGGATGGAGTTTCACTCTGTTGGCCAGCTGGAGTG 60

Qy 61 CAATGGCACAATCTCAGCTTACTGCAACCTCCGCGGTTTCAAGCGATTCTCCTGTC 120

Db	61	CAATGGCACAAATCTCAGCTTACTGCACACTTCGGCTCCGGGTTCAGCGATTCTCTCTGC	120
Qy	121	CTCAGCCTCTCAAGTAGCTGGATTACAGGCATGTGCCACACCCCTGGCTAACTAATTT	180
Db	121	CTCAGCCTCTCAAGTAGCTGGATTACAGGCATGTGCCACACCCCTGGCTAACTAATTT	180
Qy	181	CTTTTCTAATTTAGTAGAGATGGGGTTTCCATGTTGGTCAGGCTGTCTTGAATCCTCTG	240
Db	181	CTTTTCTAATTTAGTAGAGATGGGGTTTCCATGTTGGTCAGGCTGTCTTGAATCCTCTG	240
Qy	241	ACCTCAGGTGATCCACTTCGCTTGGCCCTCCCAAGTCTAGGATTCAGACCGGTGAACCTG	300
Db	241	ACCTCAGGTGATCCACTTCGCTTGGCCCTCCCAAGTCTAGGATTCAGACCGGTGAACCTG	300
Qy	301	TGCGTGGCTGATCTCTTTTGTGTTGTTGATTTTGTAAACAGGGTCTCCCTTGGTGGCC	360
Db	301	TGCGTGGCTGATCTCTTTTGTGTTGTTGATTTTGTAAACAGGGTCTCCCTTGGTGGCC	360
Qy	361	AGGCTGAGTGCAGTGTGGATCTTGGCTCACTATAACCTCCACCTCTCTGGTTCAAGT	420
Db	361	AGGCTGAGTGCAGTGTGGATCTTGGCTCACTATAACCTCCACCTCTCTGGTTCAAGT	420
Qy	421	GATCCTCCCACTTTAGCCTCTCAGTAGCTGTGATTTACAGCGGTGCACACCAACCCGG	480
Db	421	GATCCTCCCACTTTAGCCTCTCAGTAGCTGTGATTTACAGCGGTGCACACCAACCCGG	480
Qy	481	CTAATTTTGTATTTTATTATAGACACAGGGTTTCCACATGTTGGCCAGGTGTTCTCAAA	540
Db	481	CTAATTTTGTATTTTATTATAGACACAGGGTTTCCACATGTTGGCCAGGTGTTCTCAAA	540
Qy	541	CTCCTGACTCAAGGGATCCGCTCGCTCCACTTTCCCAAAGTCCCGAGATTACAGGTG	600
Db	541	CTCCTGACTCAAGGGATCCGCTCGCTCCACTTTCCCAAAGTCCCGAGATTACAGGTG	600
Qy	601	AGTCACCATGCTGACCTTATAATCTTAAAGTCAATTTTTCTGGTCCATTTCTTCCTTAG	660
Db	601	AGTCACCATGCTGACCTTATAATCTTAAAGTCAATTTTTCTGGTCCATTTCTTCCTTAG	660
Qy	661	GGTCCTCACACAAATCTGCATTAGCGGTACAAATATCCTTAACCTCAATGATTCACAAA	720
Db	661	GGTCCTCACACAAATCTGCATTAGCGGTACAAATATCCTTAACCTCAATGATTCACAAA	720
Qy	721	AGGAAGATGAAGTGATTCATGATTTAGAAAGGGGAAGTAGTAAGCCCACTGCACACTCT	780
Db	721	AGGAAGATGAAGTGATTCATGATTTAGAAAGGGGAAGTAGTAAGCCCACTGCACACTCT	780
Qy	781	GGATGATGATCCYMAATCCAGATACAGTAAAAATGGGGTAGGGAAGGTAGATACAAAA	840
Db	781	GGATGATGATCCYMAATCCAGATACAGTAAAAATGGGGTAGGGAAGGTAGATACAAAA	840
Qy	841	TTTGGTTTAAATTAATTTATCTAAATATCTAAAAACATTTTTGGATACATTTGTTGATGTA	900
Db	841	TTTGGTTTAAATTAATTTATCTAAATATCTAAAAACATTTTTGGATACATTTGTTGATGTA	900
Qy	901	ATGTAAGACTGTACAGACTCTCTAGAAAAAGTTTGGGTTCCATCTTTTCATTTCCCCCAG	960
Db	901	ATGTAAGACTGTACAGACTCTCTAGAAAAAGTTTGGGTTCCATCTTTTCATTTCCCCCAG	960
Qy	961	TGCAGTTTTCTGTAGAAATCGAATCCGAGATTTAAGTGGCAGAGAAATGACAAATGATT	1020
Db	961	TGCAGTTTTCTGTAGAAATCGAATCCGAGATTTAAGTGGCAGAGAAATGACAAATGATT	1020
Qy	1021	CCATAATGAACAAAGTCGAGAGACATTAATAATTAAGTTTAAAAATGAAGACTTACTGATG	1080
Db	1021	CCATAATGAACAAAGTCGAGAGACATTAATAATTAAGTTTAAAAATGAAGACTTACTGATG	1080
Qy	1081	AACTAAGCTTGAATAAAATTTCTGCTGATCTACAGATACTCGGNAACGTGTTAACCAAA	1140
Db	1081	AACTAAGCTTGAATAAAATTTCTGCTGATCTACAGATACTCGGNAACGTGTTAACCAAA	1140
Qy	1141	TTATGATGTGGCAAAACAAACCCAGAGACTGGTTGAGTTGTTGTTGCTCAAACTAGAGAAA	1200

QY 2281 AGTCACCAATATCAACATCTAAATGTTTGAACCCAAATCTATTGTAAGACACCA 2340
 Db 2281 AGTCACCAATATCAACATCTAAATGTTTGAACCCAAATCTATTGTAAGACACCA 2340
 QY 2341 GCAGCAATACCTGGATGATACATAGCTGTTTGAACCTCCAGTTGTAAGAAATGACT 2400
 Db 2341 GCAGCAATACCTGGATGATACATAGCTGTTTGAACCTCCAGTTGTAAGAAATGACT 2400
 QY 2401 TTCCACCTGTGTTGTCAGTGTGCAACCTTATGGCCAACTGGCTGTTTCCAGCAGCAAC 2460
 Db 2401 TTCCACCTGTGTTGTCAGTGTGCAACCTTATGGCCAACTGGCTGTTTCCAGCAGCAAC 2460
 QY 2461 AGCATCAAAATACCTGGCACTCCACTTCAAAATTTACAGGTTTACGATCTTCTTACGCAA 2520
 Db 2461 AGCATCAAAATACCTGGCACTCCACTTCAAAATTTACAGGTTTACGATCTTCTTACGCAA 2520
 QY 2521 ATGAATGCAATTCGGTTTAAAGGAGAAATTTATCCATATTAAGACAGATAGGAAGTGGAG 2580
 Db 2521 ATGAATGCAATTCGGTTTAAAGGAGAAATTTATCCATATTAAGACAGATAGGAAGTGGAG 2580
 QY 2581 GTTCAAGCAAGGATTTTCAAGTGTAAATGAAAGAAACAGATATATGCTTATAAATATG 2640
 Db 2581 GTTCAAGCAAGGATTTTCAAGTGTAAATGAAAGAAACAGATATATGCTTATAAATATG 2640
 QY 2641 TGAATTAACATCAACACAGATGATAGATCATCCGACTTATGATTTAAGAAATGCAATG 2700
 Db 2641 TGAATTAACATCAACACAGATGATAGATCATCCGACTTATGATTTAAGAAATGCAATG 2700
 QY 2701 TGAATTAACATCAACACAGATGATAGATCATCCGACTTATGATTTAAGAAATGCAATG 2760
 Db 2701 TGAATTAACATCAACACAGATGATAGATCATCCGACTTATGATTTAAGAAATGCAATG 2760
 QY 2761 ACCAGTACATCTACATGATGATGAGTGTGGAATATTTGATCTTAAATAGTGGCTTAA 2820
 Db 2761 ACCAGTACATCTACATGATGATGAGTGTGGAATATTTGATCTTAAATAGTGGCTTAA 2820
 QY 2821 AGAAAAATCCATGATCCATGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
 Db 2821 AGAAAAATCCATGATCCATGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
 QY 2881 TTCAACAATCCATCAACATGGCATTTGTTCAAGTGTCTTAAACAGCTTAACTTTCTGA 2940
 Db 2881 TTCAACAATCCATCAACATGGCATTTGTTCAAGTGTCTTAAACAGCTTAACTTTCTGA 2940
 QY 2941 TAGTTGATGGAATGCTAAAGCTAATTTGATTTTGGGATTCGAAACCAATGCAACAGATA 3000
 Db 2941 TAGTTGATGGAATGCTAAAGCTAATTTGATTTTGGGATTCGAAACCAATGCAACAGATA 3000
 QY 3001 CAACAAGTGTGTTAAAGATTTCTCAGGTTGGCAGATTAATTTATATCCACCAAGCA 3060
 Db 3001 CAACAAGTGTGTTAAAGATTTCTCAGGTTGGCAGATTAATTTATATCCACCAAGCA 3060
 QY 3061 TCAAGATATGCTCTTCCAGAGAGATGGGAATCTAAGTCAAGATAGGCCCAAAA 3120
 Db 3061 TCAAGATATGCTCTTCCAGAGAGATGGGAATCTAAGTCAAGATAGGCCCAAAA 3120
 QY 3121 GTGATGTTGGTCTTACGATGATTTTGTACTATATGACTTACGGGAAACACCAATTC 3180
 Db 3121 GTGATGTTGGTCTTACGATGATTTTGTACTATATGACTTACGGGAAACACCAATTC 3180
 QY 3181 AGCAGATTAATTAACAGATTTCTAAATTAACATGCAATTTGATCTTAAATCATGAATG 3240
 Db 3181 AGCAGATTAATTAACAGATTTCTAAATTAACATGCAATTTGATCTTAAATCATGAATG 3240
 QY 3241 AATTTCCGATATCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAGAGG 3300
 Db 3241 AATTTCCGATATCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAGAGG 3300
 QY 3301 ACCAAAAACAGAGGATATCCATCTCTGAGCTCTGGCTCATCCATATGTTCAAAATCAA 3360
 Db 3301 ACCAAAAACAGAGGATATCCATCTCTGAGCTCTGGCTCATCCATATGTTCAAAATCAA 3360

QY 3361 CTCATCCAGTTAAACAAATGGCCAAAGGAAACCACTGAAGAAATGAAATATGTTCTGGGCC 3420
 Db 3361 CTCATCCAGTTAAACAAATGGCCAAAGGAAACCACTGAAGAAATGAAATATGTTCTGGGCC 3420
 QY 3421 AACTTGTGTTGCTGAATTTCTCTAACTCCATTTTGAAGCTGCTTAAACCTTTTATATGAAC 3480
 Db 3421 AACTTGTGTTGCTGAATTTCTCTAACTCCATTTTGAAGCTGCTTAAACCTTTTATATGAAC 3480
 QY 3481 ACTATAGTGGTGAAGAGTCAATATTTCTCATCTCCAGACTTTTGAAGAAAAAGGG 3540
 Db 3481 ACTATAGTGGTGAAGAGTCAATATTTCTCATCTCCAGACTTTTGAAGAAAAAGGG 3540
 QY 3541 GAAAAAATGATTTTCAGTTATTCGTAATGTCAGATAGAGGTATAAATATATTTGGACT 3600
 Db 3541 GAAAAAATGATTTTCAGTTATTCGTAATGTCAGATAGAGGTATAAATATATTTGGACT 3600
 QY 3601 GTTATATCTCTTGAATCCCTGTGGAAATCTACATTTTGAAGCAACATCTCTGAAAGTGT 3660
 Db 3601 GTTATATCTCTTGAATCCCTGTGGAAATCTACATTTTGAAGCAACATCTCTGAAAGTGT 3660
 QY 3661 ATCAGCAAAAAAATTCAGTGAGATTTCTTTAAAGAAACTGTAAAAATAGCAACAC 3720
 Db 3661 ATCAGCAAAAAAATTCAGTGAGATTTCTTTAAAGAAACTGTAAAAATAGCAACAC 3720
 QY 3721 TTATGCACTGTATATATTTAGACTTGTCTCTGTTTCTGTTTATGCTCTTGTGTAATCTAC 3780
 Db 3721 TTATGCACTGTATATATTTAGACTTGTCTCTGTTTCTGTTTATGCTCTTGTGTAATCTAC 3780
 QY 3781 TTGACATCATTTTACTCTTGGAAATAGTGGTGGATAGCAAGTATATTTCTAAAAA 3840
 Db 3781 TTGACATCATTTTACTCTTGGAAATAGTGGTGGATAGCAAGTATATTTCTAAAAA 3840
 QY 3841 TAAATAAAGTTTGTGGCTTAAATGA 3866
 Db 3841 TAAATAAAGTTTGTGGCTTAAATGA 3866

RESULT 3

US-10-172-118-899
 ; Sequence 899, Application US/10172118
 ; Publication No. JS20030224374A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dai, Hongyue
 ; APPLICANT: He, Yudong
 ; APPLICANT: Linsley, Peter
 ; APPLICANT: Mao, Mao
 ; APPLICANT: Roberts, Chris
 ; APPLICANT: Van't Veer, Laura
 ; APPLICANT: Van de Vijver, Marc
 ; APPLICANT: Bernards, Rene
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; FILE REFERENCE: 9301-175-999
 ; CURRENT APPLICATION NUMBER: US/10/172,118
 ; PRIOR FILING DATE: 2002-06-14
 ; PRIOR APPLICATION NUMBER: 60/380,770
 ; PRIOR FILING DATE: 2002-05-14
 ; NUMBER OF SEQ ID NOS: 2699
 ; SEQ ID NO 899
 ; LENGTH: 3866
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: NM_003318
 ; DATABASE ENTRY DATE: 2001-06-18
 ; US-10-172-118-899

Query Match 100.0%; Score 3866; DB 13; Length 3866;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3866; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGAATTCCTTTTTTTTTTTTTTTTGAATGAGATTCACCTCTTGTGGCCAGGCTGGAGTG 60
 Db 1 GGAATTCCTTTTTTTTTTTTTTTTGAATGAGATTCACCTCTTGTGGCCAGGCTGGAGTG 60

QY 61 CAATGGACAATCTCAGCTTACTCAACCTCCGCTCCCGGTTCAAGCGATTCTCTGC 120
 Db 61 CAATGGACAATCTCAGCTTACTCAACCTCCGCTCCCGGTTCAAGCGATTCTCTGC 120
 QY 121 CTCAGCTCTCAAGTAGCTGGATTACAGGATGTGCCACACCCCTGGCTAACTAATTT 180
 Db 121 CTCAGCTCTCAAGTAGCTGGATTACAGGATGTGCCACACCCCTGGCTAACTAATTT 180
 QY 181 CTTTCTTATTTAGTAGAGATGGGTTTCAACATGTGCTCAGCTGTCTTGAACCTCTG 240
 Db 181 CTTTCTTATTTAGTAGAGATGGGTTTCAACATGTGCTCAGCTGTCTTGAACCTCTG 240
 QY 241 ACCTCAGGTGATCAGCTTGGCTTGGCTCCCAAGTCTAGGATTAAGCGCTGAAACTG 300
 Db 241 ACCTCAGGTGATCAGCTTGGCTTGGCTCCCAAGTCTAGGATTAAGCGCTGAAACTG 300
 QY 301 TGCCTGGCTGATCTCTTTTGTGTTGGATTTTGAACAGGGTCTCCCTTGGTGGCC 360
 Db 301 TGCCTGGCTGATCTCTTTTGTGTTGGATTTTGAACAGGGTCTCCCTTGGTGGCC 360
 QY 361 AGGCTGGAGTGCAGTGTGGGATCTTGGCTCACTATAACCTCCACCTCTCTGGTTCAAGT 420
 Db 361 AGGCTGGAGTGCAGTGTGGGATCTTGGCTCACTATAACCTCCACCTCTCTGGTTCAAGT 420
 QY 421 GATCCTCCCACTTTAGCTCTGAGTAGCTGTGATTAAGCGGTGCAACACACCCGG 480
 Db 421 GATCCTCCCACTTTAGCTCTGAGTAGCTGTGATTAAGCGGTGCAACACACCCGG 480
 QY 481 CTAATTTTGTGATTTTATTTAGACAGAGGTTTCAACATGTTGGCAGGCTGTCTCAAA 540
 Db 481 CTAATTTTGTGATTTTATTTAGACAGAGGTTTCAACATGTTGGCAGGCTGTCTCAAA 540
 QY 541 CTCCTGACTCAAGGATCCGCTTCCACTTCCCAAGTCCCGAGATTACAGGTG 600
 Db 541 CTCCTGACTCAAGGATCCGCTTCCACTTCCCAAGTCCCGAGATTACAGGTG 600
 QY 601 AGTCACATGCTGACCTTATATCTTAAGTCAATTTTCTGGTCCATTTCTTCTTAG 660
 Db 601 AGTCACATGCTGACCTTATATCTTAAGTCAATTTTCTGGTCCATTTCTTCTTAG 660
 QY 661 GGTCCTCAACAATACTGCAATAGCGGTACAAATATCTTAACTTCATGATTCAAAA 720
 Db 661 GGTCCTCAACAATACTGCAATAGCGGTACAAATATCTTAACTTCATGATTCAAAA 720
 QY 721 AGGAGATGAAGTATTCATGATTTAGAAAGGGAAGTAGTAAGCCACTGCACTCT 780
 Db 721 AGGAGATGAAGTATTCATGATTTAGAAAGGGAAGTAGTAAGCCACTGCACTCT 780
 QY 781 GGATGATGATCCCTAAATCCAGATACAGTAAATATGGGATGGGAAGGTAGAAATACAAA 840
 Db 781 GGATGATGATCCCTAAATCCAGATACAGTAAATATGGGATGGGAAGGTAGAAATACAAA 840
 QY 841 TTTGGTTTAAATTAATCTAAATATCTAAATATCTAAATATCTAAATATCTAAATATCT 900
 Db 841 TTTGGTTTAAATTAATCTAAATATCTAAATATCTAAATATCTAAATATCTAAATATCT 900
 QY 901 ATGTAAGACTGTACAGACTTCTTAGAAACAGTTTGGGTTCCATCTTTTCAATTTCCCCAG 960
 Db 901 ATGTAAGACTGTACAGACTTCTTAGAAACAGTTTGGGTTCCATCTTTTCAATTTCCCCAG 960
 QY 961 TGCAGTTTCTGTAGAAATGGAATCCGAGATTAAAGTGGCAGAGAAATGCAAAATGATT 1020
 Db 961 TGCAGTTTCTGTAGAAATGGAATCCGAGATTAAAGTGGCAGAGAAATGCAAAATGATT 1020
 QY 1021 CCATATGAACAAGTCAGAGACATTAATAATAGTTTAAATAGTTTAAATAGTTTAAATAGTT 1080
 Db 1021 CCATATGAACAAGTCAGAGACATTAATAATAGTTTAAATAGTTTAAATAGTTTAAATAGTT 1080
 QY 1081 AACTAAGCTTGAATAAATTTCTGCTGATACAGATAACTCGGGAAGTGTAAACCAA 1140
 Db 1081 AACTAAGCTTGAATAAATTTCTGCTGATACAGATAACTCGGGAAGTGTAAACCAA 1140

QY 1141 TTATGATGATGGCAAAACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAA 1200
 Db 1141 TTATGATGATGGCAAAACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAA 1200
 QY 1201 ACAGTGTTCGGTAAAGTATGCTCTTTTAAATAAATTAATTAATTAATTAATTAATTAAT 1260
 Db 1201 ACAGTGTTCGGTAAAGTATGCTCTTTTAAATAAATTAATTAATTAATTAATTAATTAAT 1260
 QY 1261 TTGAAGCCCTTCCCGCCAGATAAATATGCGCAAAATGAGAGTTTGTGAGAAATCAAGTGA 1320
 Db 1261 TTGAAGCCCTTCCCGCCAGATAAATATGCGCAAAATGAGAGTTTGTGAGAAATCAAGTGA 1320
 QY 1321 GATTTGCTGAATTAAGCTATTCAAGAGCCAGATGATGACGCTGACTTTCCTCAATGG 1380
 Db 1321 GATTTGCTGAATTAAGCTATTCAAGAGCCAGATGATGACGCTGACTTTCCTCAATGG 1380
 QY 1381 CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAATGT 1440
 Db 1381 CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAATGT 1440
 QY 1441 CACAGGTAATGTCAAAAAAAGTAAACAACTTCTTCAAAAAGCTGTAGAACGTGGACAG 1500
 Db 1441 CACAGGTAATGTCAAAAAAAGTAAACAACTTCTTCAAAAAGCTGTAGAACGTGGACAG 1500
 QY 1501 TACCACTAGAAATGCTGGAAATTTGCCCTGGGAAATTTAAACCTCCAAAAAGCAGCTGC 1560
 Db 1501 TACCACTAGAAATGCTGGAAATTTGCCCTGGGAAATTTAAACCTCCAAAAAGCAGCTGC 1560
 QY 1561 TTTTCAGAGGAGGAAAGAAATTTATCAGCATCTACGGTATTAACTTGCCTCAAGAAATCAT 1620
 Db 1561 TTTTCAGAGGAGGAAAGAAATTTATCAGCATCTACGGTATTAACTTGCCTCAAGAAATCAT 1620
 QY 1621 TTTTCGGTTTCACTTGGGCAATTTACAGAAATAGGAACAAACAGTTGTGATTCAGAGGACAGA 1680
 Db 1621 TTTTCGGTTTCACTTGGGCAATTTACAGAAATAGGAACAAACAGTTGTGATTCAGAGGACAGA 1680
 QY 1681 CTACTAAAGCCAGGTTTTTATATGAGAGAACATGCCACCAAGATGCAAGAAATAGGTT 1740
 Db 1681 CTACTAAAGCCAGGTTTTTATATGAGAGAACATGCCACCAAGATGCAAGAAATAGGTT 1740
 QY 1741 ACCGGAATTCATTGAGACAAACTAAACAACTAAACAACTAAACAACTAAACAACTAAAC 1800
 Db 1741 ACCGGAATTCATTGAGACAAACTAAACAACTAAACAACTAAACAACTAAACAACTAAAC 1800
 QY 1801 CAGTTAACTTCTAAATAGCCCAATTTGATGTGAGAGACAGATGATTCAGTTGTACCTT 1860
 Db 1801 CAGTTAACTTCTAAATAGCCCAATTTGATGTGAGAGACAGATGATTCAGTTGTACCTT 1860
 QY 1861 GTTTTATGAAAGACAAACCTCTAGATCAGAAATCCCGAGATTTGGTTGTGCTGGATCTA 1920
 Db 1861 GTTTTATGAAAGACAAACCTCTAGATCAGAAATCCCGAGATTTGGTTGTGCTGGATCTA 1920
 QY 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTTCAAAATAGTCAAT 1980
 Db 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTTCAAAATAGTCAAT 1980
 QY 1981 TCAAGGAACCTCTGCTGTGATGAAAGAGTTTCTGAACTTATTATTACTGATTCAATAA 2040
 Db 1981 TCAAGGAACCTCTGCTGTGATGAAAGAGTTTCTGAACTTATTATTACTGATTCAATAA 2040
 QY 2041 CCCTGAAAGAAATAAACCGAATCAAGTCTTCTAGCTAAATTAAGAAAGAACTAAAGAGTATC 2100
 Db 2041 CCCTGAAAGAAATAAACCGAATCAAGTCTTCTAGCTAAATTAAGAAAGAACTAAAGAGTATC 2100
 QY 2101 AAGAACACAGAGTTCCAGAGAGTAAACAGAAACAGTGGCAAGCTTAAGAGAAAGTCAGAGT 2160
 Db 2101 AAGAACACAGAGTTCCAGAGAGTAAACAGAAACAGTGGCAAGCTTAAGAGAAAGTCAGAGT 2160
 QY 2161 GTATTAAACCAAGATCTGCTGCATCTTCAAACTCACTGGGAGATTCGGGAGTTAGCCCGAA 2220
 Db 2161 GTATTAAACCAAGATCTGCTGCATCTTCAAACTCACTGGGAGATTCGGGAGTTAGCCCGAA 2220
 QY 2221 AAGTTAATACAGAGCAGAAACATACCACTTTTGGAGCAACCTGCTTTTTCAGTTTCAAAAC 2280

```
Db 2221 AAGTTAATACAGACAGAAACATACCACTTTTGGGCAACCTGCTTTTCAGTTTCAAAC 2280
Qy 2281 AGTCACACCAATATCAACATCTAATGGTTTGGCCCAAAATCTATTTGTAAGACACAA 2340
Db 2281 AGTCACACCAATATCAACATCTAATGGTTTGGCCCAAAATCTATTTGTAAGACACAA 2340
Qy 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTAAAGTCCAGTTGTAAGAAATGACT 2400
Db 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTAAAGTCCAGTTGTAAGAAATGACT 2400
Qy 2401 TTCCACCTGCTGTGCTGATGCTTAACACCTTATGGCCAACTGCTGTTTCCAGCAGCAAC 2460
Db 2401 TTCCACCTGCTGTGCTGATGCTTAACACCTTATGGCCAACTGCTGTTTCCAGCAGCAAC 2460
Qy 2461 AGCATCAAAATCTTCCACCTCCACTTCAAAATTTTACAGTTTATAGCATCTTCTTCAGCAA 2520
Db 2461 AGCATCAAAATCTTCCACCTCCACTTCAAAATTTTACAGTTTATAGCATCTTCTTCAGCAA 2520
Qy 2521 ATGAATGCATTTCCGTTTAAAGGAAGAAATTTATTCATATTTAAAGCAGATAGGAAGTGAG 2580
Db 2521 ATGAATGCATTTCCGTTTAAAGGAAGAAATTTATTCATATTTAAAGCAGATAGGAAGTGAG 2580
Qy 2581 GTTCAAGCAGGATTTTCAAGTGTAAATGAAGAAGAACAGATATATGCTATTAATATG 2640
Db 2581 GTTCAAGCAGGATTTTCAAGTGTAAATGAAGAAGAACAGATATATGCTATTAATATG 2640
Qy 2641 TGAATCTAGAAAGACAGATAACCAACTCTTGATAGTTACCGGAACGAATAGCTTATTT 2700
Db 2641 TGAATCTAGAAAGACAGATAACCAACTCTTGATAGTTACCGGAACGAATAGCTTATTT 2700
Qy 2701 TGAATAACTACAAACACAGTGAATAGATCATCCGACTTTATGATTTATGAATCAGCG 2760
Db 2701 TGAATAACTACAAACACAGTGAATAGATCATCCGACTTTATGATTTATGAATCAGCG 2760
Qy 2761 ACCAGTACATCTACATGTAATGAGTGTGAATAATTTGATCTTAAATAGTTGGCTTAAAA 2820
Db 2761 ACCAGTACATCTACATGTAATGAGTGTGAATAATTTGATCTTAAATAGTTGGCTTAAAA 2820
Qy 2821 AGAAAAATCCATTTGATCCATGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
Db 2821 AGAAAAATCCATTTGATCCATGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
Qy 2881 TTCACACATCCATCAACATGGCATTTTACAGTGATCTTAAACAGCTTAACTTTCTGA 2940
Db 2881 TTCACACATCCATCAACATGGCATTTTACAGTGATCTTAAACAGCTTAACTTTCTGA 2940
Qy 2941 TAGTTGATGGAATGCTAAAGCTAATTTGATTTGGGATTCGAAACCAATGCAACCCAGATA 3000
Db 2941 TAGTTGATGGAATGCTAAAGCTAATTTGATTTGGGATTCGAAACCAATGCAACCCAGATA 3000
Qy 3001 CAACAGAGTTGTTAAAGATTTCTCAGTTGGCAGAGTTAATATATGCGCACAGAGCAA 3060
Db 3001 CAACAGAGTTGTTAAAGATTTCTCAGTTGGCAGAGTTAATATATGCGCACAGAGCAA 3060
Qy 3061 TCAAGATATGCTTCTCCACAGAGAGATGGAATCTAAGTCAAGATAGGCCCAAAA 3120
Db 3061 TCAAGATATGCTTCTCCACAGAGAGATGGAATCTAAGTCAAGATAGGCCCAAAA 3120
Qy 3121 GTGATGTTGGTCCCTTAGATGATTTTGTACTATATGACTTACGGGAAACACCATTTTC 3180
Db 3121 GTGATGTTGGTCCCTTAGATGATTTTGTACTATATGACTTACGGGAAACACCATTTTC 3180
Qy 3181 AGCAGATAATTAATCAGATTTCTAAATACATGCGCAATTTGATCCCTAATCATGAATTC 3240
Db 3181 AGCAGATAATTAATCAGATTTCTAAATACATGCGCAATTTGATCCCTAATCATGAATTC 3240
Qy 3241 AATTTCCCGATATTCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAAAGGG 3300
Db 3241 AATTTCCCGATATTCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAAAGGG 3300
Qy 3301 ACCCAAAACAGAGATATCCATTTCTGAGCTCCCTGGCTCATCCATATGTTCAAAATCAA 3360
```

```
Db 3301 ACCCAAAACAGAGATATCCATTTCTGAGCTCCTGGCTCATCCATATGTTCAAAATCAA 3360
Qy 3361 CTCATCCAGTTAAACCAATGGCCAGGAAACCACTGAAGAATGAAATATGTTCTGGGCC 3420
Db 3361 CTCATCCAGTTAAACCAATGGCCAGGAAACCACTGAAGAATGAAATATGTTCTGGGCC 3420
Qy 3421 AACTTTGTTGGTCTGAATTTCTCTAACTCCATTTTGAAGCTGCTAAAACTTTATATGAAC 3480
Db 3421 AACTTTGTTGGTCTGAATTTCTCTAACTCCATTTTGAAGCTGCTAAAACTTTATATGAAC 3480
Qy 3481 ACTATAGTGGTGGTCAAACTATATTTCTTCTCATCTCCAGACTTTTGAAGAAAGGG 3540
Db 3481 ACTATAGTGGTGGTCAAACTATATTTCTTCTCATCTCCAGACTTTTGAAGAAAGGG 3540
Qy 3541 GAAAAAATGATTTTCAGTTTTCAGTTTTCGTAATGTCAGATAGGAGTATAAATATATTTGGACT 3600
Db 3541 GAAAAAATGATTTTCAGTTTTCAGTTTTCGTAATGTCAGATAGGAGTATAAATATATTTGGACT 3600
Qy 3601 GTTATATCTTGAATCCCTGTCGGAATCTACATTTGAAGACAACATCATCTCTGAAGTGT 3660
Db 3601 GTTATATCTTGAATCCCTGTCGGAATCTACATTTGAAGACAACATCATCTCTGAAGTGT 3660
Qy 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTTAAAGAAACTGTAAAAATAGCAACAC 3720
Db 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTTAAAGAAACTGTAAAAATAGCAACAC 3720
Qy 3721 TTATGGCACTGATATATTTAGACTTGTGTTCTCTGTTTATGCTCTTGTGTAATCTAC 3780
Db 3721 TTATGGCACTGATATATTTAGACTTGTGTTCTCTGTTTATGCTCTTGTGTAATCTAC 3780
Qy 3781 TTGACATCATTTTACTCTTGGATAGTGGGTGATAGCAAGTATATCTTAAAAAATTG 3840
Db 3781 TTGACATCATTTTACTCTTGGATAGTGGGTGATAGCAAGTATATCTTAAAAAATTG 3840
Qy 3841 TAAATAAAGTTTTTGTGGCTTAAATGA 3866
Db 3841 TAAATAAAGTTTTTGTGGCTTAAATGA 3866
```

RESULT 4

```
US-10-081-119-13
; Sequence 13, Application US/10081119
; Publication No. US20030045491A1
; GENERAL INFORMATION:
; APPLICANT: Reinhard, Christoph
; APPLICANT: Jefferson, Anne B.
; APPLICANT: Chan, Vivien W.
; TITLE OF INVENTION: TTK in Diagnosis and as a Therapeutic
; FILE REFERENCE: 16932.002
; CURRENT FILING DATE: 2002-02-21
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 3866
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1026)...(3551)
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: TTK
US-10-081-119-13
```

Query Match 100.0%; Score 3866; DB 15; Length 3866;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3866; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAATTCCTTTTTTTTTTTTTTTTGGATGAGGATTCACCTCTGTTGGCCAGGCTGGATG 60

Db 1741 ACCGGAATTCATTGAGACAAACTAAACAAACTAAACAGTCATGCCCATTTGGGAAGAGTCC 1800
Qy 1801 CAGTTAACCTTTAAATAGCCCGAGATGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Db 1801 CAGTTAACCTTTAAATAGCCCGAGATGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Qy 1861 GTTTTATGAAAAGACAAACCTCTAGATCAGAAATGCGGAGATTTGGTTGTGCGCTGGATCTA 1920
Db 1861 GTTTTATGAAAAGACAAACCTCTAGATCAGAAATGCGGAGATTTGGTTGTGCGCTGGATCTA 1920
Qy 1921 AACCAAGTGAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
Db 1921 AACCAAGTGAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
Qy 1981 TCAAGGAACCTCTGGTCTCAGATGAAAAGAGTCTCTCAACTTATTAATCTGATTCATATA 2040
Db 1981 TCAAGGAACCTCTGGTCTCAGATGAAAAGAGTCTCTCAACTTATTAATCTGATTCATATA 2040
Qy 2041 CCCTGAAGATTAACCGAATCAAGTCTTCTAGTAAATTTAGAGAACTTAAAGAGTATC 2100
Db 2041 CCCTGAAGATTAACCGAATCAAGTCTTCTAGTAAATTTAGAGAACTTAAAGAGTATC 2100
Qy 2101 AAGAACAGAGGTTCCAGAGAGTAAACAGAAACAGTGGCAAGCTTAAGAGAAAGTCAGAGT 2160
Db 2101 AAGAACAGAGGTTCCAGAGAGTAAACAGAAACAGTGGCAAGCTTAAGAGAAAGTCAGAGT 2160
Qy 2161 GTATTAAACCAAGATCTGTGTCATCTTCAATCACTGGCAGATTCGGAGTTAGCCCGAA 2220
Db 2161 GTATTAAACCAAGATCTGTGTCATCTTCAATCACTGGCAGATTCGGAGTTAGCCCGAA 2220
Qy 2221 AAGTTATACAGACGAGAAACATACCACTTTTGGACCAACCTGTCTTTTCAAGTTTCAAAAC 2280
Db 2221 AAGTTATACAGACGAGAAACATACCACTTTTGGACCAACCTGTCTTTTCAAGTTTCAAAAC 2280
Qy 2281 AGTCACCAACCAATTAACATCTAAATGGTTTGACCCCAAAATCTATTGTGAAGACACCAA 2340
Db 2281 AGTCACCAACCAATTAACATCTAAATGGTTTGACCCCAAAATCTATTGTGAAGACACCAA 2340
Qy 2341 GCACCAATACCTTGGATGATACATGAGCTGTTTGTAGAACTCCAGTTGTAAAGAAATGACT 2400
Db 2341 GCACCAATACCTTGGATGATACATGAGCTGTTTGTAGAACTCCAGTTGTAAAGAAATGACT 2400
Qy 2401 TTCACCTGCTGTGTCAGTTGTCAACACCTTATGCGCAACCTGCTGTTTCCAGCAGCAAC 2460
Db 2401 TTCACCTGCTGTGTCAGTTGTCAACACCTTATGCGCAACCTGCTGTTTCCAGCAGCAAC 2460
Qy 2461 AGCATCAAAATCTTGCCACTCCAATCAAAATTTTACAGGTTTACGATCTTCTTTCAGCAA 2520
Db 2461 AGCATCAAAATCTTGCCACTCCAATCAAAATTTTACAGGTTTACGATCTTCTTTCAGCAA 2520
Qy 2521 ATGAATGCAATTCGGTTAAAGGAAGAAATTTATCCATATTAAGCAGATAGGAAGTGGAG 2580
Db 2521 ATGAATGCAATTCGGTTAAAGGAAGAAATTTATCCATATTAAGCAGATAGGAAGTGGAG 2580
Qy 2581 GTTCAAGCAAGGATTTTCAAGTGTAAATGAAAAGAAACAGATATATGCTATAAATATG 2640
Db 2581 GTTCAAGCAAGGATTTTCAAGTGTAAATGAAAAGAAACAGATATATGCTATAAATATG 2640
Qy 2641 TGAACCTTAGAAGACAGATTAACCAACTCTTGTAGTACCGGAACGAATAGCTTATT 2700
Db 2641 TGAACCTTAGAAGACAGATTAACCAACTCTTGTAGTACCGGAACGAATAGCTTATT 2700
Qy 2701 TGAATAAACTCAACACACAGTGTATAGATCATCCGACTTTATGATTAATGAATCAGG 2760
Db 2701 TGAATAAACTCAACACACAGTGTATAGATCATCCGACTTTATGATTAATGAATCAGG 2760
Qy 2761 ACCAGTACATCTACATGATGAGTGTGAAATATTGATCTTAAATAGTTGGCTTAAAA 2820
Db 2761 ACCAGTACATCTACATGATGAGTGTGAAATATTGATCTTAAATAGTTGGCTTAAAA 2820
Qy 2821 AGAAAAATCCATTCATCCATGGGAAACCAAGAGTTACTGGAAAAATATGTTAGAGGAG 2880
Db 2821 AGAAAAATCCATTCATCCATGGGAAACCAAGAGTTACTGGAAAAATATGTTAGAGGAG 2880

Qy 2881 TTCCACAAATCCATCAACATGGCATTTGTTCACAGTGATCTTAAACCAGCTAACCTTTCTGA 2940
Db 2881 TTCCACAAATCCATCAACATGGCATTTGTTCACAGTGATCTTAAACCAGCTAACCTTTCTGA 2940
Qy 2941 TAGTTGATGGAATGCTTAAGCTTAATGATTTTGGGATTTGCAACCAATGCAACAGATA 3000
Db 2941 TAGTTGATGGAATGCTTAAGCTTAATGATTTTGGGATTTGCAACCAATGCAACAGATA 3000
Qy 3001 CAACAAGTGTGTGTTAAAGATTTCTCAGGTTGGCAGGTTAAATATATGCCCAGGAAGCAA 3060
Db 3001 CAACAAGTGTGTGTTAAAGATTTCTCAGGTTGGCAGGTTAAATATATGCCCAGGAAGCAA 3060
Qy 3061 TCAAGATATATGCTTCTCCAGAGAGATGGAATCTAAGTCAAGATAGAGTACCCCAAA 3120
Db 3061 TCAAGATATATGCTTCTCCAGAGAGATGGAATCTAAGTCAAGATAGAGTACCCCAAA 3120
Qy 3121 GTGATGTTTGGTCTTCTAGGATGATTTTGTACTATATGACTTACGGGAAAAACACCATTTT 3180
Db 3121 GTGATGTTTGGTCTTCTAGGATGATTTTGTACTATATGACTTACGGGAAAAACACCATTTT 3180
Qy 3181 AGCAGATTAATTAATCAGATTTCTAAATTAATGATGCCATAATTTGATCCTTAATCATGAAATG 3240
Db 3181 AGCAGATTAATTAATCAGATTTCTAAATTAATGATGCCATAATTTGATCCTTAATCATGAAATG 3240
Qy 3241 AATTTCCCGATATTCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGAGGG 3300
Db 3241 AATTTCCCGATATTCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGAGGG 3300
Qy 3301 ACCCAAAACAGAGGATATCCATTTCTGAGCTCTGGCTCATCCATATGTTTCAAAATCAA 3360
Db 3301 ACCCAAAACAGAGGATATCCATTTCTGAGCTCTGGCTCATCCATATGTTTCAAAATCAA 3360
Qy 3361 CTCATCCAGTTAAACCAATGGCCAAAGGAAACCACTGGAAGAAATGAAATATGTTCTGGGCC 3420
Db 3361 CTCATCCAGTTAAACCAATGGCCAAAGGAAACCACTGGAAGAAATGAAATATGTTCTGGGCC 3420
Qy 3421 AACTTGTGTTGCTGTAATTTCTCTAACTCCTTAATCTTCAATTTGAAAGCTGCTAAAACTTTATATGAAC 3480
Db 3421 AACTTGTGTTGCTGTAATTTCTCTAACTCCTTAATCTTCAATTTGAAAGCTGCTAAAACTTTATATGAAC 3480
Qy 3481 ACTATAGTGGTGGTGAAGTCAATTTCTTCACTCCAGACTTTTGAAGAAAAAGGG 3540
Db 3481 ACTATAGTGGTGGTGAAGTCAATTTCTTCACTCCAGACTTTTGAAGAAAAAGGG 3540
Qy 3541 GAAAAAATGATTTGAGTATTTGTAATGTCAGATAGGAGTATAAATATATTTGGACT 3600
Db 3541 GAAAAAATGATTTGAGTATTTGTAATGTCAGATAGGAGTATAAATATATTTGGACT 3600
Qy 3601 GTTATACCTCTTGAATCCCTGTGCAAAATCTACATTTTGAAGACAAATCATCTCTGAAGTGT 3660
Db 3601 GTTATACCTCTTGAATCCCTGTGCAAAATCTACATTTTGAAGACAAATCATCTCTGAAGTGT 3660
Qy 3661 ATCAGCAAAAAAATTCAGTGAGATTTCTTTTAAAGAAACTGTAAAAATAGCAACAC 3720
Db 3661 ATCAGCAAAAAAATTCAGTGAGATTTCTTTTAAAGAAACTGTAAAAATAGCAACAC 3720
Qy 3721 TTATGGCACTGTATATATTTAGACTTGTCTCTGTTTATGCTTGTGTAATCTAC 3780
Db 3721 TTATGGCACTGTATATATTTAGACTTGTCTCTGTTTATGCTTGTGTAATCTAC 3780
Qy 3781 TTGACATCATTTTACTCTTTGGAAATAGTGGTGGATAGCAAGTATATTTTAAAAAACTTTG 3840
Db 3781 TTGACATCATTTTACTCTTTGGAAATAGTGGTGGATAGCAAGTATATTTTAAAAAACTTTG 3840
Qy 3841 TAAATAAAGTTTGTGGCTAAAAATGA 3866
Db 3841 TAAATAAAGTTTGTGGCTAAAAATGA 3866

RESULT 6
US-10-116-712-663
; Sequence 663, Application US/10116712

Db 1861 GTTTTATGAAAGAGCAAACTCTAGATCAGATGCGAGATTTGGTGTGCGTGGATCTA 1920
Qy 1921 AACCAAGTGAAGATGATCTCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
Db 1921 AACCAAGTGAAGATGATCTCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
Qy 1981 TCAGGAACTCTGTGTGAGATGAAAGAGTCTGAACCTTATTATTACTGATTCAATAA 2040
Db 1981 TCAAGGAACCTCTGTGTGAGATGAAAGAGTCTGAACCTTATTATTACTGATTCAATAA 2040
Qy 2041 CCTGGAAGATAAAGCGAATCAAGTCTTCTAGCTAAATTAAGAAATCAATAAGATATC 2100
Db 2041 CCTGGAAGATAAAGCGAATCAAGTCTTCTAGCTAAATTAAGAAATCAATAAGATATC 2100
Qy 2101 AAGNACAGAGGTTCCAGAGAGTAAACAGAAACAGTGGCAAGCTAAGAGAAAGTCAAGT 2160
Db 2101 AAGNACAGAGGTTCCAGAGAGTAAACAGAAACAGTGGCAAGCTAAGAGAAAGTCAAGT 2160
Qy 2161 GTATTAACCAAGATCCCTGCTGCATCTTCAAACTCACTGGCAGATTCGGGAGTTAGCCCGAA 2220
Db 2161 GTATTAACCAAGATCCCTGCTGCATCTTCAAACTCACTGGCAGATTCGGGAGTTAGCCCGAA 2220
Qy 2221 AAGTTAATACAGAGCAGAAACATACACTTTTGAAGCAACCTGTCTTTTCAAGTTTCAAAC 2280
Db 2221 AAGTTAATACAGAGCAGAAACATACACTTTTGAAGCAACCTGTCTTTTCAAGTTTCAAAC 2280
Qy 2281 AGTCACCAACCAATCAACATCTAAATGTTTGAACCCAAATCTATTGTTGAAGACACCAA 2340
Db 2281 AGTCACCAACCAATCAACATCTAAATGTTTGAACCCAAATCTATTGTTGAAGACACCAA 2340
Qy 2341 GCAGCAATACCTTGGATGATTAACAGAGCTGTTTGAAGCTCAAGTGTGTAAGAAATGACT 2400
Db 2341 GCAGCAATACCTTGGATGATTAACAGAGCTGTTTGAAGCTCAAGTGTGTAAGAAATGACT 2400
Qy 2401 TTCCACCTGCTGTGTCAGTGTGTCACACCTTATGSCCAACCTGCTGTTTCCAGCAGCAAC 2460
Db 2401 TTCCACCTGCTGTGTCAGTGTGTCACACCTTATGSCCAACCTGCTGTTTCCAGCAGCAAC 2460
Qy 2461 AGCATCAAAATACCTTGCCACTCCACTTCAAAATTTACAGGTTTGTAGCATCTTCTTCAGCAA 2520
Db 2461 AGCATCAAAATACCTTGCCACTCCACTTCAAAATTTACAGGTTTGTAGCATCTTCTTCAGCAA 2520
Qy 2521 ATGAATGCAATTCGGTTAAAGGAAGATTTATTCATATTAAGACAGATAGGAGTGGAG 2580
Db 2521 ATGAATGCAATTCGGTTAAAGGAAGATTTATTCATATTAAGACAGATAGGAGTGGAG 2580
Qy 2581 GTTCAAGCAAGGTATTTCAAGTGTAAATGAAAGAAACAGATATATGCTATATAAATATG 2640
Db 2581 GTTCAAGCAAGGTATTTCAAGTGTAAATGAAAGAAACAGATATATGCTATATAAATATG 2640
Qy 2641 TGAACCTTAGAAGAGCAGATAACCAACTCTTGTAGTATCCGGAACGAAATAGCTTATT 2700
Db 2641 TGAACCTTAGAAGAGCAGATAACCAACTCTTGTAGTATCCGGAACGAAATAGCTTATT 2700
Qy 2701 TGAATTAACCAACACACAGTGTATGATCATCCGACTTTATGATTAAGAAATCAGG 2760
Db 2701 TGAATTAACCAACACACAGTGTATGATCATCCGACTTTATGATTAAGAAATCAGG 2760
Qy 2761 ACCAGTACATCTACATGTAATGAGTGTGAAATATTGATCTTAAATAGTTGGCTTAAA 2820
Db 2761 ACCAGTACATCTACATGTAATGAGTGTGAAATATTGATCTTAAATAGTTGGCTTAAA 2820
Qy 2821 AGAAAAATCCATTCATCGGGAACCGCAGAGTTACTGGAAAAATATGTTAGGCGAG 2880
Db 2821 AGAAAAATCCATTCATCGGGAACCGCAGAGTTACTGGAAAAATATGTTAGGCGAG 2880
Qy 2881 TTCAACACATCCATCAACATGGCATTTTCAAGTGTCTTAAACCCAGCTAACCTTCTGA 2940
Db 2881 TTCAACACATCCATCAACATGGCATTTTCAAGTGTCTTAAACCCAGCTAACCTTCTGA 2940
Qy 2941 TAGTTGATGGAATGCTAAAGCTAATGATTTTGGGATTTGGAATGGAACCAATGCAACAGATA 3000

Db 2941 TAGTTGATGGAATGCTAAAGCTAATTTGATTTTGGGATTTGCAACCAAAATGCAACCGAGATA 3000
Qy 3001 CAACAGAGTGTGTTAAAGATTTCTCAGGTTGCGCAGATTAATTTATATGCGCACCAAGAGCAA 3060
Db 3001 CAACAGAGTGTGTTAAAGATTTCTCAGGTTGCGCAGATTAATTTATATGCGCACCAAGAGCAA 3060
Qy 3061 TCAAGATATATGTTCTTCCAGAGAGAAATGGAATCTAAGTCAAGATAGAGCCCAAAA 3120
Db 3061 TCAAGATATATGTTCTTCCAGAGAGAAATGGAATCTAAGTCAAGATAGAGCCCAAAA 3120
Qy 3121 GTGATGTTTGGTCTTAGGATGTTTGTACTATATGACTTACGGGAAAAACACCATTTTC 3180
Db 3121 GTGATGTTTGGTCTTAGGATGTTTGTACTATATGACTTACGGGAAAAACACCATTTTC 3180
Qy 3181 AGCAGATATTAATCAGATTTCTAAATTTACATGCGCATTAATTTGATCTTAATCATGAATTTG 3240
Db 3181 AGCAGATATTAATCAGATTTCTAAATTTACATGCGCATTAATTTGATCTTAATCATGAATTTG 3240
Qy 3241 AATTTCCCGATATTTCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAGAGGG 3300
Db 3241 AATTTCCCGATATTTCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAGAGGG 3300
Qy 3301 ACCCAAAAACAGAGGATATCCATTTCTGAGCTCTCTGGCTCATCCATATGTTCAAAATTTCAA 3360
Db 3301 ACCCAAAAACAGAGGATATCCATTTCTGAGCTCTCTGGCTCATCCATATGTTCAAAATTTCAA 3360
Qy 3361 CTCATCCAGTTAACCAATGCGCAAGGAAACCACTGAAGAAATGAATATGTTCTGGGCCC 3420
Db 3361 CTCATCCAGTTAACCAATGCGCAAGGAAACCACTGAAGAAATGAATATGTTCTGGGCCC 3420
Qy 3421 AACTTTGTTGGTCTGAATTTCTCCTAATCTCCTAATCTCAATTTTGAAGCTGCTAATAATGAAC 3480
Db 3421 AACTTTGTTGGTCTGAATTTCTCCTAATCTCCTAATCTCAATTTTGAAGCTGCTAATAATGAAC 3480
Qy 3481 ACTATAGTGGTGGTGAAGTCAATAATCTTCAATCTCCAGACTTTTGAAAAAAAAGGG 3540
Db 3481 ACTATAGTGGTGGTGAAGTCAATAATCTTCAATCTCCAGACTTTTGAAAAAAAAGGG 3540
Qy 3541 GAAAAAATGATTTTCAGTATTTTCGATGTCAGATAGGAGTATAAATATATTTGGACT 3600
Db 3541 GAAAAAATGATTTTCAGTATTTTCGATGTCAGATAGGAGTATAAATATATTTGGACT 3600
Qy 3601 GTTATACTCTTGAATCCCTGTGGAATCTTACATTTTGAAGACAAATCATCTCTGAAAGTGT 3660
Db 3601 GTTATACTCTTGAATCCCTGTGGAATCTTACATTTTGAAGACAAATCATCTCTGAAAGTGT 3660
Qy 3661 ATCAGCAAAAAAATTTCAAGTGAATATCTTTTAAAGAAACTGTAAAAATAGCAACAC 3720
Db 3661 ATCAGCAAAAAAATTTCAAGTGAATATCTTTTAAAGAAACTGTAAAAATAGCAACAC 3720
Qy 3721 TTATGSCACTGTATATATTGTAGACTTGTCTTCTGTTTATGCTCTTGTGTAATCTAC 3780
Db 3721 TTATGSCACTGTATATATTGTAGACTTGTCTTCTGTTTATGCTCTTGTGTAATCTAC 3780
Qy 3781 TTGACATCAATTTTACTCTTTGGAATAGTGGTGGATAGCAAGTATATTTTAAAAAATTTG 3840
Db 3781 TTGACATCAATTTTACTCTTTGGAATAGTGGTGGATAGCAAGTATATTTTAAAAAATTTG 3840
Qy 3841 TAAATTAAGTGTGCTGCTTAAATGA 3866
Db 3841 TAAATTAAGTGTGCTGCTTAAATGA 3866

RESULT 7

US-10-353-690-1
; Sequence 1, Application US/10353690
; Publication No. US20030215840A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Chun, Miyoung
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.

Qy	1321	GATTGCTGAATTAAAAAGCTATTCAAGAGCCAGATGATGCACGTGACTACTTTCAAATGG	1380
Db	1321	GATTGCTGAATTAAAAAGCTATTCAAGAGCCAGATGATGCACGTGACTACTTTCAAATGG	1380
Qy	1381	CCAGAGCAAACTGCAAGAAAATTTGCTTTTGTTCATATATCTTTTGTGACAAATTTTGAACGTG	1440
Db	1381	CCAGAGCAAACTGCAAGAAAATTTGCTTTTGTTCATATATCTTTTGTGACAAATTTTGAACGTG	1440
Qy	1441	CACAAGGTAATGTCAAAAAAAGTAAACAACTTCTTCAAAAAGCTGTAGAACGTGGAGCAG	1500
Db	1441	CACAAGGTAATGTCAAAAAAAGTAAACAACTTCTTCAAAAAGCTGTAGAACGTGGAGCAG	1500
Qy	1501	TACCACTAGAAAATGCTGGAAAATTTGCCCTGCGGAAATTTAAACCTTCCAAAAAAGACAGCTGC	1560
Db	1501	TACCACTAGAAAATGCTGGAAAATTTGCCCTGCGGAAATTTAAACCTTCCAAAAAAGACAGCTGC	1560
Qy	1561	TTTCAGAGGAGGAAAAGAAATTTATCAGCATCTACGGTATTAACTGCCCCCAAGAATCAT	1620
Db	1561	TTTCAGAGGAGGAAAAGAAATTTATCAGCATCTACGGTATTAACTGCCCCCAAGAATCAT	1620
Qy	1621	TTTCCGGTTCACTTGGGCATTTACAGAAATAGGAACAAACAGTTGTGATTCAGAGGACAGA	1680
Db	1621	TTTCCGGTTCACTTGGGCATTTACAGAAATAGGAACAAACAGTTGTGATTCAGAGGACAGA	1680
Qy	1681	CTACTAAGCCAGGTTTTTATATGAGAGAACATGCCACACAAAGATGCAGAAATAGGTT	1740
Db	1681	CTACTAAGCCAGGTTTTTATATGAGAGAACATGCCACACAAAGATGCAGAAATAGGTT	1740
Qy	1741	ACCGGAATTCATTGAGACAAACTAACAAAACTAAACAGTCATGCCCAATTTGGGAAGTGCC	1800
Db	1741	ACCGGAATTCATTGAGACAAACTAACAAAACTAAACAGTCATGCCCAATTTGGGAAGTGCC	1800
Qy	1801	CAGTTAACCTCTAAATAGCCAGATTTGATGTGAAGACAGATGATTCAGTTGTACTCTT	1860
Db	1801	CAGTTAACCTCTAAATAGCCAGATTTGATGTGAAGACAGATGATTCAGTTGTACTCTT	1860
Qy	1861	GTTTTATGAAAAGACAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCCTGGATCTA	1920
Db	1861	GTTTTATGAAAAGACAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCCTGGATCTA	1920
Qy	1921	AACCAAGTGGAATGATTCCTGTGAATTAGAAATTTAAAGTCTGTTTCAAAATAGTCAATT	1980
Db	1921	AACCAAGTGGAATGATTCCTGTGAATTAGAAATTTAAAGTCTGTTTCAAAATAGTCAATT	1980
Qy	1981	TCAAGGAACCTCTGGTGTGATGAAAGAGTTCTGAACTATTATTACTGATTCAATAA	2040
Db	1981	TCAAGGAACCTCTGGTGTGATGAAAGAGTTCTGAACTATTATTACTGATTCAATAA	2040
Qy	2041	CCCTGAAGAAATAAACCGAATCAAGCTTCTAGCTAAATTAGAAGAAACTAAAGAGTATC	2100
Db	2041	CCCTGAAGAAATAAACCGAATCAAGCTTCTAGCTAAATTAGAAGAAACTAAAGAGTATC	2100
Qy	2101	AAGAACACAGAGTTCCAGAGAGTAACACAGAAACAGTCGCAGCTTAACAGAAAGTCAGAGT	2160
Db	2101	AAGAACACAGAGTTCCAGAGAGTAACACAGAAACAGTCGCAGCTTAACAGAAAGTCAGAGT	2160
Qy	2161	GTATTAAACCGAATCCTGCTGCATCTTTCAAATCACTGGCAGATTTCCGGAGTTAGCCCGAA	2220
Db	2161	GTATTAAACCGAATCCTGCTGCATCTTTCAAATCACTGGCAGATTTCCGGAGTTAGCCCGAA	2220
Qy	2221	AAGTTAATACAGACAGAAACATACCACCTTTTGTAGCAACCTGCTCTTTTCAGTTTCAAAC	2280
Db	2221	AAGTTAATACAGACAGAAACATACCACCTTTTGTAGCAACCTGCTCTTTTCAGTTTCAAAC	2280
Qy	2281	AGTCAACACCAATATCAACATCTAAATGGTTTGAACCCAAAATCTATTGTGAAGACACAA	2340
Db	2281	AGTCAACACCAATATCAACATCTAAATGGTTTGAACCCAAAATCTATTGTGAAGACACAA	2340
Qy	2341	GCAGCAATACCTTGGATGATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAAGAAATGACT	2400
Db	2341	GCAGCAATACCTTGGATGATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAAGAAATGACT	2400

Qy	2401	TTCCACCTGCTGTGTCAGTTGCTAACACCTTATATGGCCAACTGCGCTGTGTTTCCAGCAGCAAC	2460
Db	2401	TTCCACCTGCTGTGTCAGTTGCTAACACCTTATATGGCCAACTGCGCTGTGTTTCCAGCAGCAAC	2460
Qy	2461	AGCATCAAACTACTTTGCCACTCCACTTCACAAATTTTACAGGTTTTAGCATCTTCTTCAGCAA	2520
Db	2461	AGCATCAAACTACTTTGCCACTCCACTTCACAAATTTTACAGGTTTTAGCATCTTCTTCAGCAA	2520
Qy	2521	ATGAATGCATTTTCGGTTTAAAGGAAGAAATTTATTCCTATTTAAAGCGAGATAGGAAGTGGAG	2580
Db	2521	ATGAATGCATTTTCGGTTTAAAGGAAGAAATTTATTCCTATTTAAAGCGAGATAGGAAGTGGAG	2580
Qy	2581	GTTTCAAGCAAGGTATTTTCAGGTGTTAAATCAAAGAAAACAGATATATGCTATATAAAATATG	2640
Db	2581	GTTTCAAGCAAGGTATTTTCAGGTGTTAAATCAAAGAAAACAGATATATGCTATATAAAATATG	2640
Qy	2641	TGAACCTTAGAAGACAGATAACCAAACTCTTGATAGTTACCGGAAACGAATAGCTTATTT	2700
Db	2641	TGAACCTTAGAAGACAGATAACCAAACTCTTGATAGTTACCGGAAACGAATAGCTTATTT	2700
Qy	2701	TGAATAAATCTACACACACACAGTGATTAAGATCATCCGACTTTTATGATTATGAATACACGG	2760
Db	2701	TGAATAAATCTACACACACACAGTGATTAAGATCATCCGACTTTTATGATTATGAATACACGG	2760
Qy	2761	ACCAGTACATCTCATCGTGAATGAGGTGTGGAATATTTGATCTTAAATAGTTGGCTTAAAA	2820
Db	2761	ACCAGTACATCTCATCGTGAATGAGGTGTGGAATATTTGATCTTAAATAGTTGGCTTAAAA	2820
Qy	2821	AGAAAAATCCATTTGATCCATGGGAACGCAAGAGTTTACTGCGAAAAATATGTTAGAGGCAG	2880
Db	2821	AGAAAAATCCATTTGATCCATGGGAACGCAAGAGTTTACTGCGAAAAATATGTTAGAGGCAG	2880
Qy	2881	TTCCACACAATCCATCAACATGGCAATTCCTCACAGTGTATCTTAAACACAGCTTAACTTCTGA	2940
Db	2881	TTCCACACAATCCATCAACATGGCAATTCCTCACAGTGTATCTTAAACACAGCTTAACTTCTGA	2940
Qy	2941	TAGTTGATGGAAATGCTTAAAGCTAATTCGATTTTGGGATTCGAAACCAAAATGCAACACGATA	3000
Db	2941	TAGTTGATGGAAATGCTTAAAGCTAATTCGATTTTGGGATTCGAAACCAAAATGCAACACGATA	3000
Qy	3001	CAACAAGTGTGTTTAAAGATTTCTCAGGTTGCGCACAGTTAATTTATGCCACCAAGAACAA	3060
Db	3001	CAACAAGTGTGTTTAAAGATTTCTCAGGTTGCGCACAGTTAATTTATGCCACCAAGAACAA	3060
Qy	3061	TCAAAGATATGCTTCTCTCCAGAGAGAAATGGGAAATCTAAGTCAAAGATAAGCCCCAAAA	3120
Db	3061	TCAAAGATATGCTTCTCTCCAGAGAGAAATGGGAAATCTAAGTCAAAGATAAGCCCCAAAA	3120
Qy	3121	GTGATGTTTGGTCTCTTAGATGATTTTGTGACTATATGACTTACGGGAAAACACCATTTTC	3180
Db	3121	GTGATGTTTGGTCTCTTAGATGATTTTGTGACTATATGACTTACGGGAAAACACCATTTTC	3180
Qy	3181	AGCAGATAAATTAATCAGATTTCTAAATTTACATGCCATAATTTGATCTCTTAAATCATGAAATTG	3240
Db	3181	AGCAGATAAATTAATCAGATTTCTAAATTTACATGCCATAATTTGATCTCTTAAATCATGAAATTG	3240
Qy	3241	AATTTCCGGATATTCCAGAGAAAGATCTTCAAGATGTGTTTAAAGTGTGTTTAAAAAGGG	3300
Db	3241	AATTTCCGGATATTCCAGAGAAAGATCTTCAAGATGTGTTTAAAGTGTGTTTAAAAAGGG	3300
Qy	3301	ACCCAAAACACAGGATATCCATTCTCAGCTCCCTGGCTCATCATATGTTTCAAAATTCAAA	3360
Db	3301	ACCCAAAACACAGGATATCCATTCTCAGCTCCCTGGCTCATCATATGTTTCAAAATTCAAA	3360
Qy	3361	CTCATCCAGTTTAAACCAATGCCAAGGGAACCACTGAAGAAATGAAATATGTTCTCGGCC	3420
Db	3361	CTCATCCAGTTTAAACCAATGCCAAGGGAACCACTGAAGAAATGAAATATGTTCTCGGCC	3420
Qy	3421	AACTTGTGTGCTGAATTTCTCCTAACTCCATTTTGAAGCTGCTTAAACCTTTATATGAAC	3480
Db	3421	AACTTGTGTGCTGAATTTCTCCTAACTCCATTTTGAAGCTGCTTAAACCTTTATATGAAC	3480
Qy	3481	ACTATAGTGTGGTGAAGGTCTAATTTCTTCATCTCTCCAGACCTTTTGAAAAAAAAAGGG	3540

Db 3481 ACTATAGTGGTGAAGTCAATAATCTTCATCCTCAAGACTTTTGAAGAAAAAGGG 3540
 Qy 3541 GAAAAAATGATTTGCAGTATTCGTAATGTCAGATAGGAGGTATAAATATATTGACT 3600
 Db 3541 GAAAAAATGATTTGCAGTATTCGTAATGTCAGATAGGAGGTATAAATATATTGACT 3600
 Qy 3601 GTTATCTCTTGAATCCCTGTGGAATCTCAATTTGAAGACAACATCACTCTGAAGTGT 3660
 Db 3601 GTTATCTCTTGAATCCCTGTGGAATCTCAATTTGAAGACAACATCACTCTGAAGTGT 3660
 Qy 3661 ATCAGCAAAAAAATCAGTGAATATCTTTAAAGAAAACTGTAAATATAGCAACAC 3720
 Db 3661 ATCAGCAAAAAAATCAGTGAATATCTTTAAAGAAAACTGTAAATATAGCAACAC 3720
 Qy 3721 TTATGGCACGTATATATTGTAGACTTGTCTCTCTTTATGCTCTTGTGTAATCTAC 3780
 Db 3721 TTATGGCACGTATATATTGTAGACTTGTCTCTCTTTATGCTCTTGTGTAATCTAC 3780
 Qy 3781 TTGACATCAATTTACTCTTGAATAGTGGTGGATAGCAAGTATATCTTAAAAAATTG 3840
 Db 3781 TTGACATCAATTTACTCTTGAATAGTGGTGGATAGCAAGTATATCTTAAAAAATTG 3840
 Qy 3841 TAAATAAAGTTTCTGGCTAAATGA 3866
 Db 3841 TAAATAAAGTTTCTGGCTAAATGA 3866

RESULT 9

US-10-295-027-803
 ; Sequence 803, Application US/10295027
 ; Publication No US2003023250A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Afar, Daniel
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Ginsberg, Wendy M.
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Hevezi, Peter A.
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Watson, Susan R.
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
 ; FILE REFERENCE: 018501-012500US
 ; CURRENT APPLICATION NUMBER: US/10/295,027
 ; CURRENT FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: US 09/663,733
 ; PRIOR FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/335,394
 ; PRIOR FILING DATE: 2001-11-15
 ; PRIOR APPLICATION NUMBER: US 60/332,464
 ; PRIOR FILING DATE: 2001-11-21
 ; PRIOR APPLICATION NUMBER: US 60/334,393
 ; PRIOR FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: US 60/340,376
 ; PRIOR FILING DATE: 2001-12-14
 ; PRIOR APPLICATION NUMBER: US 60/347,211
 ; PRIOR FILING DATE: 2002-01-08
 ; PRIOR APPLICATION NUMBER: US 60/347,349
 ; PRIOR FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 60/355,250
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: US 60/356,714
 ; PRIOR FILING DATE: 2002-02-13
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 1386
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 803
 ; LENGTH: 3866

TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-295-027-803
 Query Match 100.0%; Score 3866; DB 16; Length 3866;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3866; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GGAATTCCTTTTTTTTTTTTTTTTTTTTGGAGTGGAGTTTCACTCTTGTGGCAGGCTGGAGTG 60
 Db 1 GGAATTCCTTTTTTTTTTTTTTTTTTTTGGAGTGGAGTTTCACTCTTGTGGCAGGCTGGAGTG 60
 Qy 61 CAATGGCAAACTCTCAGCTTACTGCAACTCCGGCTCCCGGTTCAAGCGATTCCTCTGC 120
 Db 61 CAATGGCAAACTCTCAGCTTACTGCAACTCCGGCTCCCGGTTCAAGCGATTCCTCTGC 120
 Qy 121 CTCAGCCTCTCAAGTAGCTGGATACAGGATGTGCCACCACTCCCTGGCTAACTAATTT 180
 Db 121 CTCAGCCTCTCAAGTAGCTGGATACAGGATGTGCCACCACTCCCTGGCTAACTAATTT 180
 Qy 181 CTTTCTATTTAGTAGAGTGGGTTTCAACATGTGTCTCAGGCTGGTCTTGAATCTCTG 240
 Db 181 CTTTCTATTTAGTAGAGTGGGTTTCAACATGTGTCTCAGGCTGGTCTTGAATCTCTG 240
 Qy 241 ACCTCAGGTGATCCACTTGGCTTGGCTCCCAAGTGTAGGATACAGCGTGAACCTG 300
 Db 241 ACCTCAGGTGATCCACTTGGCTTGGCTCCCAAGTGTAGGATACAGCGTGAACCTG 300
 Qy 301 TGCCTGGCTGATCTCTTTTTTTTGGATTGTTGAAACAGGCTCTCCCTTGGTCCGCC 360
 Db 301 TGCCTGGCTGATCTCTTTTTTTTGGATTGTTGAAACAGGCTCTCCCTTGGTCCGCC 360
 Qy 361 AGGTGGAGTGCAGTGGTGGATCTTGGCTCAGTATTAACCTCCACCTCTCTGGTTTCAAGT 420
 Db 361 AGGTGGAGTGCAGTGGTGGATCTTGGCTCAGTATTAACCTCCACCTCTCTGGTTTCAAGT 420
 Qy 421 GATCTCTCCACTTTAGCTCTGAGTGTGTGATACAGCGTGCACCACTCCACCCGG 480
 Db 421 GATCTCTCCACTTTAGCTCTGAGTGTGTGATACAGCGTGCACCACTCCACCCGG 480
 Qy 481 CTAATTTTGTATTTTATTAGACAGAGGTTTCAACATGTGTGATACAGCGTGCACCACT 540
 Db 481 CTAATTTTGTATTTTATTAGACAGAGGTTTCAACATGTGTGATACAGCGTGCACCACT 540
 Qy 541 CTCCTGGACTCAAGGATCCGCTCCCTCCACTTCCCAAGTCCCGAGATACAGGTTG 600
 Db 541 CTCCTGGACTCAAGGATCCGCTCCCTCCACTTCCCAAGTCCCGAGATACAGGTTG 600
 Qy 601 AGTCACCATGCTGACCTTAAATTTCTTAAGTCAATTTTTTCTGGTCCACTTCTTCTTAG 660
 Db 601 AGTCACCATGCTGACCTTAAATTTCTTAAGTCAATTTTTTCTGGTCCACTTCTTCTTAG 660
 Qy 661 GGTCTCTACACAAATCTGCATTTAGGCGGTACAAATATCTTAATCTTCACTGATTCACAAA 720
 Db 661 GGTCTCTACACAAATCTGCATTTAGGCGGTACAAATATCTTAATCTTCACTGATTCACAAA 720
 Qy 721 AGGAAGATGAAGTGAATTCATGATTTAGAAAGGGAAGTAGTAAAGCCCACTGCACTCTCT 780
 Db 721 AGGAAGATGAAGTGAATTCATGATTTAGAAAGGGAAGTAGTAAAGCCCACTGCACTCTCT 780
 Qy 781 GGATGATGATCTTAAATCCAGATACAGTAAATGGGATGGGAAGGTAGATACAAAA 840
 Db 781 GGATGATGATCTTAAATCCAGATACAGTAAATGGGATGGGAAGGTAGATACAAAA 840
 Qy 841 TTTGGTTTAAATTAATTAATCTAAATATCTAAATATCTAAATATCTAAATATCTTGTGATG 900
 Db 841 TTTGGTTTAAATTAATTAATCTAAATATCTAAATATCTAAATATCTTGTGATGATG 900
 Qy 901 ATGTAAGACTGTACAGACTTCTTAGAAAAACAGTTTGGGTTCCATCTTTTCCATTTCCCCAG 960
 Db 901 ATGTAAGACTGTACAGACTTCTTAGAAAAACAGTTTGGGTTCCATCTTTTCCATTTCCCCAG 960
 Qy 961 TGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGATTCACAATTGATT 1020

Qy	661	GGTCTCTCAACAAATCTGCATTAGCGGTACAAATATCCTTAATCTTATGATTCACAAA	720	1741	ACCGAATTCATTGAGACAAAATAACAAAACCTAAACAGTCAATGCCATTTCGAAGAGTCC	1800
Db	661	GGTCTCTCAACAAATCTGCATTAGCGGTACAAATATCCTTAATCTTATGATTCACAAA	720	1741	ACCGAATTCATTGAGACAAAATAACAAAACCTAAACAGTCAATGCCATTTCGAAGAGTCC	1800
Qy	721	AGGAGATGAAGTCAATTCATGATTAGAAAGGGAGTAGTAAGCCCATCGCACACTCT	780	1801	CAGTTAACTTCTAAATAGCCACAGATTGTGATGTGAAGACAGATGATTCAGTTGACCTT	1860
Db	721	AGGAGATGAAGTCAATTCATGATTAGAAAGGGAGTAGTAAGCCCATCGCACACTCT	780	1801	CAGTTAACTTCTAAATAGCCACAGATTGTGATGTGAAGACAGATGATTCAGTTGACCTT	1860
Qy	781	GGATGATGATCCCTAAATCCAGATACAGTAAATGGGTATGGGAAGGTAGATAACAAA	840	1861	GTCTTATGAAAAGACAAACCTCTAGATCAGAAATCCGAGATTGGTTGTGCTGGATCTA	1920
Db	781	GGATGATGATCCCTAAATCCAGATACAGTAAATGGGTATGGGAAGGTAGATAACAAA	840	1861	GTCTTATGAAAAGACAAACCTCTAGATCAGAAATCCGAGATTGGTTGTGCTGGATCTA	1920
Qy	841	TTTGGTTTAAATTAATTAATCTAAATATCTAAAAACATTTTGGATACATTTGATGTGA	900	1921	AACCAAGTGGAAATGATTTCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCAAT	1980
Db	841	TTTGGTTTAAATTAATTAATCTAAATATCTAAAAACATTTTGGATACATTTGATGTGA	900	1921	AACCAAGTGGAAATGATTTCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCAAT	1980
Qy	901	ATGTAAGACTGTACAGACTTCTTAGAAAAACAGTTTGGTTCCATCTTTTCAATTTCCCCAG	960	1981	TCAGGAACCTCTGCTGTGCAGATGAAAAGAGTTCTGAACCTTATTACTGATCAATAA	2040
Db	901	ATGTAAGACTGTACAGACTTCTTAGAAAAACAGTTTGGTTCCATCTTTTCAATTTCCCCAG	960	1981	TCAGGAACCTCTGCTGTGCAGATGAAAAGAGTTCTGAACCTTATTACTGATCAATAA	2040
Qy	961	TGCAGTTTCTGTAGAAATGGAATCCGAGATTTAAGTGCAGAGAAATTGACAATGAT	1020	2041	CCCTGAAGAATAAAACCGAATCAAGTCTTCTAGCTAAATTAAGAAGAACTAAAGAGTATC	2100
Db	961	TGCAGTTTCTGTAGAAATGGAATCCGAGATTTAAGTGCAGAGAAATTGACAATGAT	1020	2041	CCCTGAAGAATAAAACCGAATCAAGTCTTCTAGCTAAATTAAGAAGAACTAAAGAGTATC	2100
Qy	1021	CCATAATGAAAGTCAAGAGACATTAATAATTAAGTTTAAATTAAGTCTTACTGATG	1080	2101	AAGAACCCAGAGTTCAGAGAGTAGTAAACAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT	2160
Db	1021	CCATAATGAAAGTCAAGAGACATTAATAATTAAGTTTAAATTAAGTCTTACTGATG	1080	2101	AAGAACCCAGAGTTCAGAGAGTAGTAAACAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT	2160
Qy	1081	AACCTAAGCTTGAATAAATTTCTGCTGATACAGATACTCGGAACTGTTTAAACAAA	1140	2161	GTATTAACAGAAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCCGGAGTTAGCCGAA	2220
Db	1081	AACCTAAGCTTGAATAAATTTCTGCTGATACAGATACTCGGAACTGTTTAAACAAA	1140	2161	GTATTAACAGAAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCCGGAGTTAGCCGAA	2220
Qy	1141	TTATGATGATGGCAACACCGAGAGACTGGTTGATGTTGTTGCTCAAACTAGAGAAA	1200	2221	AGTTTAAATACAGAGCAGAAAACATACCCTTTTGAACAACCTGCTTTTCAGTTTCAAAAC	2280
Db	1141	TTATGATGATGGCAACACCGAGAGACTGGTTGATGTTGTTGCTCAAACTAGAGAAA	1200	2221	AGTTTAAATACAGAGCAGAAAACATACCCTTTTGAACAACCTGCTTTTCAGTTTCAAAAC	2280
Qy	1201	ACAGTGTCCGCTAAGTGATGCTTTTAAATTAAGTTGATGTTGCTCAAGCAAGCAA	1260	2281	AGTCAACCAATATCAACATCTAAATCGTTTGACCCCAAAATCTATTGTGAAGACCAA	2340
Db	1201	ACAGTGTCCGCTAAGTGATGCTTTTAAATTAAGTTGATGTTGCTCAAGCAAGCAA	1260	2281	AGTCAACCAATATCAACATCTAAATCGTTTGACCCCAAAATCTATTGTGAAGACCAA	2340
Qy	1261	TTGAAGCGCTTCCCCAGATAAATATGCGCAAAATGAGAGTTTGTGTAAGTTCAAGTGA	1320	2341	GCAGCAATACCTTGGATGATTACAGCTGTTTGTAGAACTCCAGTTCTTAAGAGATGACT	2400
Db	1261	TTGAAGCGCTTCCCCAGATAAATATGCGCAAAATGAGAGTTTGTGTAAGTTCAAGTGA	1320	2341	GCAGCAATACCTTGGATGATTACAGCTGTTTGTAGAACTCCAGTTCTTAAGAGATGACT	2400
Qy	1321	GATTTGCTGAATTAAGTATTAACAGCCAGATGATGCAGTGACTCTTTCAAAATGG	1380	2401	TTCCACCTGCTTGTGCTGTTGCAACACCTTATGGCAACCTGCTGTTTCCAGACGCAAC	2460
Db	1321	GATTTGCTGAATTAAGTATTAACAGCCAGATGATGCAGTGACTCTTTCAAAATGG	1380	2401	TTCCACCTGCTTGTGCTGTTGCAACACCTTATGGCAACCTGCTGTTTCCAGACGCAAC	2460
Qy	1381	CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCAACAATTTGAACTGT	1440	2461	AGCATCAAAATACCTTCCCATCCACTTCAAAATTTTACAGGTTTACGATCTTCTTCCAGCAA	2520
Db	1381	CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCAACAATTTGAACTGT	1440	2461	AGCATCAAAATACCTTCCCATCCACTTCAAAATTTTACAGGTTTACGATCTTCTTCCAGCAA	2520
Qy	1441	CACAAGGTATGTCAAAAAAGTAAACAACTTCTTCAAAAAGCTGTAGAACGTGGAGCAG	1500	2521	ATGAATGCATTTCCGTTTAAAGGAAGAAATTTATTCATATTAAGCAGATAGGAAGTGGAG	2580
Db	1441	CACAAGGTATGTCAAAAAAGTAAACAACTTCTTCAAAAAGCTGTAGAACGTGGAGCAG	1500	2521	ATGAATGCATTTCCGTTTAAAGGAAGAAATTTATTCATATTAAGCAGATAGGAAGTGGAG	2580
Qy	1501	TACCACCTAGAAATGCTGGAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC	1560	2581	GTCTTAAGCAAGGTATTTTCCAGTGTAAATGAAGAAAACAGATATATGCTATAAAATATG	2640
Db	1501	TACCACCTAGAAATGCTGGAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC	1560	2581	GTCTTAAGCAAGGTATTTTCCAGTGTAAATGAAGAAAACAGATATATGCTATAAAATATG	2640
Qy	1561	TTTCAGAGGAGAAAAGAGAAATTTATCAGCATCTACGGTATTAATCTGCCCCAAGATCAT	1620	2641	TGAATCTTAAGAAGCAGATACCAAACTCTTGATAGTTTACCGGAACCAAAATAGCTTATT	2700
Db	1561	TTTCAGAGGAGAAAAGAGAAATTTATCAGCATCTACGGTATTAATCTGCCCCAAGATCAT	1620	2641	TGAATCTTAAGAAGCAGATACCAAACTCTTGATAGTTTACCGGAACCAAAATAGCTTATT	2700
Qy	1621	TTTCCGGTTCACTTGGGCAATTTACAGAAATAGGAACAACTGTTGATTTCCAGAGGACAGA	1680	2701	TGAATAAACTACAACAACACAGTGTAGATCATCCGACTTTATGATTTGAAAATCACGG	2760
Db	1621	TTTCCGGTTCACTTGGGCAATTTACAGAAATAGGAACAACTGTTGATTTCCAGAGGACAGA	1680	2701	TGAATAAACTACAACAACACAGTGTAGATCATCCGACTTTATGATTTGAAAATCACGG	2760
Qy	1681	CTACTAAAGCCAGGTTTTTATATGAGAGAACATGCCACCAAGATGCAAGAAATAGGTT	1740	2761	ACAGTACATCTACATGTTAGTGTGGAATATTTGATCTTAAATAGTTGGCTTAAA	2820
Db	1681	CTACTAAAGCCAGGTTTTTATATGAGAGAACATGCCACCAAGATGCAAGAAATAGGTT	1740	2761	ACAGTACATCTACATGTTAGTGTGGAATATTTGATCTTAAATAGTTGGCTTAAA	2820
				2821	AGAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG	2880

Db 301 TGCTGGCTGATTCCTTTTGTGTTGATTTTGAACAGGGTCTCCCTTGGTCGCC 360
 QY 361 AGCTGGAGTGCAGTGGTGGCTCTGCTCACTATAAAGTCCACCTCTGTTTCAAGT 420
 Db 361 AGCTGGAGTGCAGTGGTGGCTCTGCTCACTATAAAGTCCACCTCTGTTTCAAGT 420
 QY 421 GATCCTCCCACTTTAGCCTCCTGAGTGTGATTTACAGGCGTGCACCAACACCCGG 480
 Db 421 GATCCTCCCACTTTAGCCTCCTGAGTGTGATTTACAGGCGTGCACCAACACCCGG 480
 QY 481 CTAAATTTTGTATTTTATAGAGACAGGTTTCAACATGTTGGCCAGGCTGTTCTCAA 540
 Db 481 CTAAATTTTGTATTTTATAGAGACAGGTTTCAACATGTTGGCCAGGCTGTTCTCAA 540
 QY 541 CTCTGACATCAAGGATCCGCTGCTCCACCTCCCAAGTCCCGAGATTACAGGTGTG 600
 Db 541 CTCTGACATCAAGGATCCGCTGCTCCACCTCCCAAGTCCCGAGATTACAGGTGTG 600
 QY 601 AGTCACATGCTGACCTTAAATTTTAAAGTCAATTTTCTGCTGATTTCTTCCTTAG 660
 Db 601 AGTCACATGCTGACCTTAAATTTTAAAGTCAATTTTCTGCTGATTTCTTCCTTAG 660
 QY 661 GGTCTCACAAACAACTCTGATAGGCGGTACAAATATCCCTTAACTTCATGATTCACAAA 720
 Db 661 GGTCTCACAAACAACTCTGATAGGCGGTACAAATATCCCTTAACTTCATGATTCACAAA 720
 QY 721 AGGAAGATGAAGTGAATCATGATTTAGAAAGGGGAAGTAGTAAGCCCACTGCACTCCT 780
 Db 721 AGGAAGATGAAGTGAATCATGATTTAGAAAGGGGAAGTAGTAAGCCCACTGCACTCCT 780
 QY 781 GGATGATGATCCCTAAATCCAGATACAGTAAGATGGGTATGGAGAGTGAATACAAA 840
 Db 781 GGATGATGATCCCTAAATCCAGATACAGTAAGATGGGTATGGAGAGTGAATACAAA 840
 QY 841 TTTGGTTTAAATTAATTAATCTAAATATCTAAATATCTAAATATCTAAATATCTAAAT 900
 Db 841 TTTGGTTTAAATTAATTAATCTAAATATCTAAATATCTAAATATCTAAATATCTAAAT 900
 QY 901 ATGTAAGATGTAGACATCTCTAGAAAACAGTTTGGGTTCCATCTTTTCATTTCCCCAG 960
 Db 901 ATGTAAGATGTAGACATCTCTAGAAAACAGTTTGGGTTCCATCTTTTCATTTCCCCAG 960
 QY 961 TGCAGTTTCTGTAGAAATGGAAATCCAGAGTTTAAAGTGAAGATGACAAATGAT 1020
 Db 961 TGCAGTTTCTGTAGAAATGGAAATCCAGAGTTTAAAGTGAAGATGACAAATGAT 1020
 QY 1021 CCATAATGAACAAAGTGAAGACATTTAAATTAAGTTTAAATTAAGTTTAAATTAAGTT 1080
 Db 1021 CCATAATGAACAAAGTGAAGACATTTAAATTAAGTTTAAATTAAGTTTAAATTAAGTT 1080
 QY 1081 AACTAGCTTGAATTAATTTCTGCTGATCTACAGATTAAGTGGGAACTGTTAAACCAA 1140
 Db 1081 AACTAGCTTGAATTAATTTCTGCTGATCTACAGATTAAGTGGGAACTGTTAAACCAA 1140
 QY 1141 TTATGATGATGGCAACCAACAGGAGCTGTTTGGTTGTTGCTCAAACTAGAGAAA 1200
 Db 1141 TTATGATGATGGCAACCAACAGGAGCTGTTTGGTTGTTGCTCAAACTAGAGAAA 1200
 QY 1201 ACAGTGTCCGTAAAGTGTGCTCTTTAAATTAATTTGATTTGCTGATCTACAGTCA 1260
 Db 1201 ACAGTGTCCGTAAAGTGTGCTCTTTAAATTAATTTGATTTGCTGATCTACAGTCA 1260
 QY 1261 TTGAAGCGCTTTCCCGAGATAATATGGCCAAATGAGAGTTTGGCTAGAAATCAAGTGA 1320
 Db 1261 TTGAAGCGCTTTCCCGAGATAATATGGCCAAATGAGAGTTTGGCTAGAAATCAAGTGA 1320
 QY 1321 GATTGTGTAATTAAGCTATTCAGAGCCAGATGATGACGTGACTACTTTCAATGG 1380
 Db 1321 GATTGTGTAATTAAGCTATTCAGAGCCAGATGATGACGTGACTACTTTCAATGG 1380
 QY 1381 CCAGAGCAACTGCAAGAAATTTGCTTTGTTTCATATCTTTTGGCAATTTGCACTGT 1440

Db 1381 CCAGAGCAACTGCAAGAAATTTGCTTTGTTTCATATCTTTTGCACAAATTTGAAGTGT 1440
 QY 1441 CACAGGTAATGTCAAAAAAGTAAACAACTTTCTTCAAAAAGCTGTAGAACGTGGAGCAG 1500
 Db 1441 CACAGGTAATGTCAAAAAAGTAAACAACTTTCTTCAAAAAGCTGTAGAACGTGGAGCAG 1500
 QY 1501 TACCACTAGAAATGCTGGAAATTTGCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC 1560
 Db 1501 TACCACTAGAAATGCTGGAAATTTGCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC 1560
 QY 1561 TTTCAGAGGAGGAAAAAGAAATTTATCAGCATCTACGGTATTAACTTGCCTCAAGAAATCAT 1620
 Db 1561 TTTCAGAGGAGGAAAAAGAAATTTATCAGCATCTACGGTATTAACTTGCCTCAAGAAATCAT 1620
 QY 1621 TTTCCGTTCTACTGGGCAATTTACAGATAGGAACAACAGTTGTGATTTCCAGAGACAGA 1680
 Db 1621 TTTCCGTTCTACTGGGCAATTTACAGATAGGAACAACAGTTGTGATTTCCAGAGACAGA 1680
 QY 1681 CTACTAAAGCCAGGTTTTTATATGAGAGAAACATGCCCCACCAAGATGCAAAAATAGTGT 1740
 Db 1681 CTACTAAAGCCAGGTTTTTATATGAGAGAAACATGCCCCACCAAGATGCAAAAATAGTGT 1740
 QY 1741 ACCGGAATTCATTTAGACAAAACTAAACAACTAAACAGTATGCCCATTTTGGAGAGTCC 1800
 Db 1741 ACCGGAATTCATTTAGACAAAACTAAACAACTAAACAGTATGCCCATTTTGGAGAGTCC 1800
 QY 1801 CAGTTTAAACCTTCTAAATAGCCAGATTTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
 Db 1801 CAGTTTAAACCTTCTAAATAGCCAGATTTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
 QY 1861 GTTTTATGAAAGACAAACCTCTAGATCAGAAATGCCAGATTTGGTTGTGCTGGATCTA 1920
 Db 1861 GTTTTATGAAAGACAAACCTCTAGATCAGAAATGCCAGATTTGGTTGTGCTGGATCTA 1920
 QY 1921 AACCAAGTGGAAATGATTTCTGTGAAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCAAT 1980
 Db 1921 AACCAAGTGGAAATGATTTCTGTGAAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCAAT 1980
 QY 1981 TCAAGGAACCTCTGCTGTGAGATGAAAGAGTTTCTGAACTTATTTACTGATTTCAATAA 2040
 Db 1981 TCAAGGAACCTCTGCTGTGAGATGAAAGAGTTTCTGAACTTATTTACTGATTTCAATAA 2040
 QY 2041 CCTTGAAGATTAACCGGAATCAAGTCTTTAGCTTAAATTAAGAAATTAAGAAATTAAG 2100
 Db 2041 CCTTGAAGATTAACCGGAATCAAGTCTTTAGCTTAAATTAAGAAATTAAGAAATTAAG 2100
 QY 2101 AAGAACCCAGAGTTCCAGAGATTAACCAAGTCCAGAAACAGTGGCAAGTAAAGAGTCA 2160
 Db 2101 AAGAACCCAGAGTTCCAGAGATTAACCAAGTCCAGAAACAGTGGCAAGTAAAGAGTCA 2160
 QY 2161 GTATTAAACCAAGTCTGCTGATCTTTCAAAATCACTGGCAGATTCGGAGTTAGCCGAA 2220
 Db 2161 GTATTAAACCAAGTCTGCTGATCTTTCAAAATCACTGGCAGATTCGGAGTTAGCCGAA 2220
 QY 2221 AAGTTAATACAGAGCAAGAAATCACTTTTGGCAACCTGCTTTTTCAGTTTCAAAAC 2280
 Db 2221 AAGTTAATACAGAGCAAGAAATCACTTTTGGCAACCTGCTTTTTCAGTTTCAAAAC 2280
 QY 2281 AGTCACCAATATCAACATCTAAATGTTTGGCCCAAAATCTATTTTGAAGACACCAA 2340
 Db 2281 AGTCACCAATATCAACATCTAAATGTTTGGCCCAAAATCTATTTTGAAGACACCAA 2340
 QY 2341 GGAGCAATACCTTGGATGATTTACATGAGCTGTTTTAGAACCTCCAGTTGTAAAGATGACT 2400
 Db 2341 GGAGCAATACCTTGGATGATTTACATGAGCTGTTTTAGAACCTCCAGTTGTAAAGATGACT 2400
 QY 2401 TTCCACCTCTGCTGATTTGTCAACCTTATGGCCAACTGCTGTTTCCAGAGCAAC 2460
 Db 2401 TTCCACCTCTGCTGATTTGTCAACCTTATGGCCAACTGCTGTTTCCAGAGCAAC 2460
 QY 2461 AGCATCAAAATCTGCCACTCCACTTCAAAATTTACAGGTTTTAGCATCTTCTTCAAGCA 2520
 Db 2461 AGCATCAAAATCTGCCACTCCACTTCAAAATTTACAGGTTTTAGCATCTTCTTCAAGCA 2520

QY 2521 ATGAATGCAATTCGGTTAAAGGAAGAAATTTATTTCCATATTAAGCAGATAGGAAGTGGAG 2580
DB 2521 ATGAATGCAATTCGGTTAAAGGAAGAAATTTATTTCCATATTAAGCAGATAGGAAGTGGAG 2580
QY 2581 GTTCAAGCAGAGTATTTTCAGGTGTTAAATGAAGAAGAAACAGATATATGCTATATAAATATG 2640
DB 2581 GTTCAAGCAGAGTATTTTCAGGTGTTAAATGAAGAAGAAACAGATATATGCTATATAAATATG 2640
QY 2641 TGAATTTAGAGAGCAGATAACCAAACTCTTGATAGTTACCGGAACGAAATAGCTTAT 2700
DB 2641 TGAATTTAGAGAGCAGATAACCAAACTCTTGATAGTTACCGGAACGAAATAGCTTAT 2700
QY 2701 TGAATAAAGTACAAACAGATGATGAATCATCCGACTTTATGATATGAAATCAGCG 2760
DB 2701 TGAATAAAGTACAAACAGATGATGAATCATCCGACTTTATGATATGAAATCAGCG 2760
QY 2761 ACCAGTACATCTCATGTTAATGGAGTGTGGAATATTTGATCTTAATAGTTGGCTTAAAA 2820
DB 2761 ACCAGTACATCTCATGTTAATGGAGTGTGGAATATTTGATCTTAATAGTTGGCTTAAAA 2820
QY 2821 AGAAAAATCCATGTATCCATGGGAAGCAAGAGTTACTGGAATAATATGTTAGAGCGAG 2880
DB 2821 AGAAAAATCCATGTATCCATGGGAAGCAAGAGTTACTGGAATAATATGTTAGAGCGAG 2880
QY 2881 TTCAACAATCCATCAACATGGCATTTTTCACAGTGTATTAACAGCTTAATTTCTGCA 2940
DB 2881 TTCAACAATCCATCAACATGGCATTTTTCACAGTGTATTAACAGCTTAATTTCTGCA 2940
QY 2941 TAGTTGATGAATGCTAAAGCTAAATTAATTTGGGATTCGAAACCAAAATGCAACAGATA 3000
DB 2941 TAGTTGATGAATGCTAAAGCTAAATTAATTTGGGATTCGAAACCAAAATGCAACAGATA 3000
QY 3001 CACACAGTGTGTTAAAGATCTCAGGTGGCAGATTAATATATGACCAAGCAAGCA 3060
DB 3001 CACACAGTGTGTTAAAGATCTCAGGTGGCAGATTAATATATGACCAAGCAAGCA 3060
QY 3061 TCAAGATATGTTCTCTCCAGAGAAATGGGAAATCTAAGTCAAAAGATAAGCCCAAAA 3120
DB 3061 TCAAGATATGTTCTCTCCAGAGAAATGGGAAATCTAAGTCAAAAGATAAGCCCAAAA 3120
QY 3121 GTGATGTTGGTCTTAGAGATGTTTGTGATATATGATCTTAAGCAAGCAAGCAAGCA 3180
DB 3121 GTGATGTTGGTCTTAGAGATGTTTGTGATATATGATCTTAAGCAAGCAAGCAAGCA 3180
QY 3181 AGCAGATAATTAATCAGATTTCTAAATTTACATGCCATAATGATCTCAATCATGAAATG 3240
DB 3181 AGCAGATAATTAATCAGATTTCTAAATTTACATGCCATAATGATCTCAATCATGAAATG 3240
QY 3241 AATTTCCGATATCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAGAGG 3300
DB 3241 AATTTCCGATATCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAGAGG 3300
QY 3301 ACCCAAAACAGAGATATCCATCTCGAGTCTCTGCTCATCATATGTTCAAAATCAAA 3360
DB 3301 ACCCAAAACAGAGATATCCATCTCGAGTCTCTGCTCATCATATGTTCAAAATCAAA 3360
QY 3361 CTATCCAGTTAACCAAAATGGCCAGGGAACCACTGAAGAAATGAATATGTTCTGGGCC 3420
DB 3361 CTATCCAGTTAACCAAAATGGCCAGGGAACCACTGAAGAAATGAATATGTTCTGGGCC 3420
QY 3421 AACTTCTGGTCTGATATCTCTCACTCTCACTCTCACTCTCACTCTCACTCTCACTCTCA 3480
DB 3421 AACTTCTGGTCTGATATCTCTCACTCTCACTCTCACTCTCACTCTCACTCTCACTCTCA 3480
QY 3481 ACTATAGTGGTGAAGATCATATTTCTTCATCTCCCAAGACTTTTCAAAAAAAGGG 3540
DB 3481 ACTATAGTGGTGAAGATCATATTTCTTCATCTCCCAAGACTTTTCAAAAAAAGGG 3540
QY 3541 GAAAAAATGATTTGAGTATTTGTAATGTCAGATAGGAGGTATATAATATTTGGACT 3600
DB 3541 GAAAAAATGATTTGAGTATTTGTAATGTCAGATAGGAGGTATATAATATTTGGACT 3600

QY 3601 GTTATACTTTGAATCCCTGTGGAATCTCATTTTGAAGCAACATCCTCTGAAGTGT 3660
DB 3601 GTTATACTTTGAATCCCTGTGGAATCTCATTTTGAAGCAACATCCTCTGAAGTGT 3660
QY 3661 ATCAGCAAAAAAATTCAGTGAGATTTCTTTAAAGAAAACTGTAATAATAGCAAC 3720
DB 3661 ATCAGCAAAAAAATTCAGTGAGATTTCTTTAAAGAAAACTGTAATAATAGCAAC 3720
QY 3721 TTATGGCACTGTATATATTTAGACTTTCTTTCTGTTTATCTCTTGTGTAATCTAC 3780
DB 3721 TTATGGCACTGTATATATTTAGACTTTCTTTCTGTTTATCTCTTGTGTAATCTAC 3780
QY 3781 TTGACATCAATTTTACTCTTGGATAGTGGGTGGATAGCAATATATTTCTAAAAA 3840
DB 3781 TTGACATCAATTTTACTCTTGGATAGTGGGTGGATAGCAATATATTTCTAAAAA 3840
QY 3841 TAAATAAAGTTTCTGGCTAAATGA 3866
DB 3841 TAAATAAAGTTTCTGGCTAAATGA 3866

RESULT 12

US-10-295-027-481
; Sequence 481, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 481
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-295-027-481

Query Match 66.5%; Score 2569.2; DB 16; Length 2574;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2571; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 978 ATGGAATCCGAGGATTTAAAGTGGCAGAGAAATGCAATTCATTCCATAATGCAACAAAGTG 1037
Db 1 ATGGAATCCGAGGATTTAAAGTGGCAGAGAAATGCAATTCATTCCATAATGCAACAAAGTG 60
QY 1038 AGGACATTTAAATAAGTTTAAATAAGAGACCTTACTGATGAATTAAGCTTGAATAA 1097
Db 61 AGGACATTTAAATAAGTTTAAATAAGAGACCTTACTGATGAATTAAGCTTGAATAA 120
QY 1098 ATTTCTGCTGATCTACAGATAAATCGGGAATCTGTTAAACCAATTAATGATGAGGCAAC 1157
Db 121 ATTTCTGCTGATCTACAGATAAATCGGGAATCTGTTAAACCAATTAATGATGAGGCAAC 180
QY 1158 AACCCAGAGACTGGTTGAGTTTGTCTCAAACTAGAGAAAACAGTGTTCGCTAAGT 1217
Db 181 AACCCAGAGACTGGTTGAGTTTGTCTCAAACTAGAGAAAACAGTGTTCGCTAAGT 240
QY 1218 GATGCTCTTTTAAATAAATGATTTGTTGCTTACAGTCAAGCAATGAAGGCTTCCCCCA 1277
Db 241 GATGCTCTTTTAAATAAATGATTTGTTGCTTACAGTCAAGCAATGAAGGCTTCCCCCA 300
QY 1278 GATTAATATGCCCCAAAATGAGATTTTGTCTAGAAATCAAGTGAGATTTGCTGAATTAATA 1337
Db 301 GATTAATATGCCCCAAAATGAGATTTTGTCTAGAAATCAAGTGAGATTTGCTGAATTAATA 360
QY 1338 GCTATTCAAGAGCCAGATGATGACGCTGACTTCTTCAAAATGGCCAGCAAACTGCAAG 1397
Db 361 GCTATTCAAGAGCCAGATGATGACGCTGACTTCTTCAAAATGGCCAGCAAACTGCAAG 420
QY 1398 AAATTTGCTTTTGTTCATATATCTTTTGGCAATTTGAATGTGCACAAAGTAATGCAAA 1457
Db 421 AAATTTGCTTTTGTTCATATCTTTTGGCAATTTGAATGTGCACAAAGTAATGCAAA 480
QY 1458 AAAAGTAAACAACTCTTCAAAAGCTGTAGACGTGAGACGTACCACTAGAAATGCTG 1517
Db 481 AAAAGTAAACAACTCTTCAAAAGCTGTAGACGTGAGACGTACCACTAGAAATGCTG 540
QY 1518 GAAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGTTTTCAGAGGAGGAAAAG 1577
Db 541 GAAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGTTTTCAGAGGAGGAAAAG 600
QY 1578 AAGAAATTTACAGATCTAGGTAATTAATCTGCCCAAGAAATCATTTTCGGTTCACTTGGG 1637
Db 601 AAGAAATTTACAGATCTAGGTAATTAATCTGCCCAAGAAATCATTTTCGGTTCACTTGGG 660
QY 1638 CATTTACAGATAGGAAACACAGTTGTGATTCAGAGGACAGACTACTAAAGCCAGGTTT 1697
Db 661 CATTTACAGATAGGAAACACAGTTGTGATTCAGAGGACAGACTACTAAAGCCAGGTTT 720
QY 1698 TTATATGGAGAAACATGCCCAACAGATGCAAGAAATAGGTTACCGGAATTCATTGAGA 1757
Db 721 TTATATGGAGAAACATGCCCAACAGATGCAAGAAATAGGTTACCGGAATTCATTGAGA 780
QY 1758 CAARCTAACAACTTAAACAGTCATGCCATTTTGGAGAGTCCAGTTTACCTCTTAAT 1817
Db 781 CAARCTAACAACTTAAACAGTCATGCCATTTTGGAGAGTCCAGTTTACCTCTTAAT 840
QY 1818 AGCCAGATTTGATGTGAAGACAGATGATTCAGTTGATCTGTTTATGAAAAGACAA 1877
Db 841 AGCCAGATTTGATGTGAAGACAGATGATTCAGTTGATCTGTTTATGAAAAGACAA 900
QY 1878 ACCTCTAGATCAGAAATCCGAGATTTGGTTGTGCTGGATCTAAACCAAGTGAATGAT 1937
Db 901 ACCTCTAGATCAGAAATCCGAGATTTGGTTGTGCTGGATCTAAACCAAGTGAATGAT 960
QY 1938 TCCTGTGAATTAAGAAATTTAAAGCTGTTCAAAATAGTCATTTCAAGGAACCTCTGGTG 1997
Db 961 TCCTGTGNAATTAAGAAATTTAAAGCTGTTCAAAATAGTCATTTCAAGGAACCTCTGGTG 1020
QY 1998 TCAGATGAAGAGTCTGAACTTATTAATCTGATTAATCAATCAACCTGGAAGATAAAG 2057
Db 1021 TCAGATGAAGAGTCTGAACTTATTAATCTGATTAATCAATCAACCTGGAAGATAAAG 1080

QY 2058 GAATCAAGTCTTCTAGCTAAATTTAGAGAAACTAAAGAGTATCAAGAACCGAGGTTCCA 2117
Db 1081 GAATCAAGTCTTCTAGCTAAATTTAGAGAAACTAAAGAGTATCAAGAACCGAGGTTCCA 1140
QY 2118 GAGAGTAAACAGAAACAGTGGCAAGCTTAAGAGAAAGTCAAGAGTGTATTAAACAGAAATCCT 2177
Db 1141 GAGAGTAAACAGAAACAGTGGCAATCTAAGAGAAAGTCAAGAGTGTATTAAACAGAAATCCT 1200
QY 2178 GCTGATCTTCAAACTCACTGGCAGATTCGGAGTTAGCCCGAAAAAGTTTAATACAGAGAG 2237
Db 1201 GCTGATCTTCAAACTCACTGGCAGATTCGGAGTTAGCCCGAAAAAGTTTAATACAGAGAG 1260
QY 2238 AAACATACCACTTTTGGCAACCTGTCTTTTCAAGTTTCAAAACAGTCAACCAATATCA 2297
Db 1261 AAACATACCACTTTTGGCAACCTGTCTTTTCAAGTTTCAAAACAGTCAACCAATATCA 1320
QY 2298 ACATCTAAATGTTTGAACCCAAAATCTATTGTTAAGACACCAAGCAGCAATACCTTGGAT 2357
Db 1321 ACATCTAAATGTTTGAACCCAAAATCTATTGTTAAGACACCAAGCAGCAATACCTTGGAT 1380
QY 2358 GATTACATGAGCTGTTTGAACCTCCAGTGTAAAGAAATGACTTTCCACCTGCTTGTGAG 2417
Db 1381 GATTACATGAGCTGTTTGAACCTCCAGTGTAAAGAAATGACTTTCCACCTGCTTGTGAG 1440
QY 2418 TTGTCAACACCTTTATGGCCAACTGCTGTTTCCAGCAGCAACAGCAATCAATCTGCC 2477
Db 1441 TTGTCAACACCTTTATGGCCAACTGCTGTTTCCAGCAGCAACAGCAATCAATCTGCC 1500
QY 2478 ACTCCATTTCAAAATTTACAGTGTTCAGTATCTTCTTCCAGCAAAATGAATGATTCGGTT 2537
Db 1501 ACTCCATTTCAAAATTTACAGTGTTCAGTATCTTCTTCCAGCAAAATGAATGATTCGGTT 1560
QY 2538 AAAGGAAGAAATTTATTTCCATATTAAGCAGATAGGAAGTGGAGTTTCAAGCAAGTATTT 2597
Db 1561 AAAGGAAGAAATTTATTTCCATTTAAAGCAGATAGGAAGTGGAGTTTCAAGCAAGTATTT 1620
QY 2598 CAGGTGTTAAATGAAAGAAACAGATATATGCTATATAAATATGTGAACTTTAGAAGACA 2657
Db 1621 CAGGTGTTAAATGAAAGAAACAGATATATGCTATATAAATATGTGAACTTTAGAAGACA 1680
QY 2658 GATAACCAACTCTTGTAGTTTACCGAACGAAATAGCTTATTTGAAATATACTACAAAC 2717
Db 1681 GATAACCAACTCTTGTAGTTTACCGAACGAAATAGCTTATTTGAAATATACTACAAAC 1740
QY 2718 CACAGTGAATGATCATCCGACTTTATGATTTATGAAATCAGCGACAGTACATCTACATG 2777
Db 1741 CACAGTGAATGATCATCCGACTTTATGATTTATGAAATCAGCGACAGTACATCTACATG 1800
QY 2778 GTAATGGAGTGTGGAAATATTGATCTTAATAGTTGGCTTAAAGAAAATCCATGAT 2837
Db 1801 GTAATGGAGTGTGGAAATATTGATCTTAATAGTTGGCTTAAAGAAAATCCATGAT 1860
QY 2838 CCATGGGAACGCAAGGTTTACTGGAAAAATATGTTAGAGGAGTTTCAACCAATCCATCAA 2897
Db 1861 CCATGGGAACGCAAGGTTTACTGGAAAAATATGTTAGAGGAGTTTCAACCAATCCATCAA 1920
QY 2898 CATGCAATGTTTCAAGTGAATCTTAAACCAAGCTTAACTTTCTGATAGTTGATGAAATGCTA 2957
Db 1921 CATGCAATGTTTCAAGTGAATCTTAAACCAAGCTTAACTTTCTGATAGTTGATGAAATGCTA 1980
QY 2958 AAGCTAATTTGATTTTGGGATTCGAAACCAATGCAACCAAGTACCAACAGTGTCTTAAA 3017
Db 1981 AAGCTAATTTGATTTTGGGATTCGAAACCAATGCAACCAAGTACCAACAGTGTCTTAAA 2040
QY 3018 GATTTCTCAGGTTGGCAGAGTTTAAATTTATATGCCACAGAAAGCAATCAAGATATGCTTCC 3077
Db 2041 GATTTCTCAGGTTGGCAGAGTTTAAATTTATATGCCACAGAAAGCAATCAAGATATGCTTCC 2100
QY 3078 TCCAGAGAAATGGAAATCTTAAGTCAAGATAAGCCCAAGTGAATGTTGGTCCCTA 3137
Db 2101 TCCAGAGAAATGGAAATCTTAAGTCAAGATAAGCCCAAGTGAATGTTGGTCCCTA 2160
QY 3138 GGATGATTTTGTCTATATGACTTTACGGGAAAAACCAATTTTACAGAGATAATTAATCAG 3197

Db 642 GGACAGAAGAGTGTGTCAGCATCGACAGTACTAAGTGCCCAAGAGCGGTTCTCCAGCTC 701
QY 1631 ACTTGGGCAATTACAGAAATAGGAACAACAGTGTGTGATTCAGAGGACAGACTACTAAAGC 1690
Db 702 ACTTGGAAATGTACAGAAATCGGAGCATCAGTGTGAGTCCAGAGGACAGGCTCGGGCAGC 761
QY 1691 CAGGTTTTTATATCGAGAGAAATCGCCACACAGATGCGAGAAATAGTTTACCGGAATC 1750
Db 762 CAGGTTTTTATATCGAGAGAAATCGCCACACAGATGCGGAACTGAGGCATCAAAACCC 821
QY 1751 ATTGAGACAATTAACAAAATTAACAGTCAATGCTATGCTGCTGCTGCTGCTGCTGCTGCT 1810
Db 822 CTTCAAGCAGACTCAGCAGCTTAACCGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 881
QY 1811 TCTAAATAGCCAGATGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTGTTTATGAA 1870
Db 882 TCTAAACAGCCAGATTTCTATGTGAAGACAGATGATTCAGTTGTGATGCTGCTGCTGCTGCT 941
QY 1871 AAGCAAACTCTAGATCAGAAATCCGAGATTTGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1930
Db 942 AAGAAAGGGTCCAGGACCGA---CCGAGACGGGATTCGCGCGGCTCCAGACCAAGTGG 998
QY 1931 AAATGATTCCTGTAATTAAGAAATTAAGTCTGTTCAAAATAGTCAATTCAGGAAC 1990
Db 999 CAGTGAATTCCTATGAATCGAGGTTTAAAGCCCATTCAAACTATCTATTTGAAGACTC 1058
QY 1991 TCTGGTGTGATGAAAGAGTCTGAACTTATTAATTAATTAATTAATTAATTAATTAATTAAT 2050
Db 1059 TTTGGTGTCCAATGAAAGAGTCTGAACTTATTAATTAATTAATTAATTAATTAATTAATTAAT 1115
QY 2051 TAAACCGAATCAAGTCTTCTAGCTAAATTAAGAAAGTCTAATTAATTAATTAATTAATTAAT 2110
Db 1116 TAAACAGATTAAG---TCTAACAATAATTTGAAGAACTA-----AGCCAGA 1160
QY 2111 GGTTCAGAGAGTAACAGAAACAGTGGCAAGCTTAAGAGAAAGTCAAGTGTATTAACCA 2170
Db 1161 GATTCAGAGAAAGGCGCCATGAGTGGAGTCTACCAAGAAAGCGGAGTGTGTTCCA 1220
QY 2171 GAATCTCTGTCATCTTCAATCACTGGGAGATTCGGAGTTAGCCGGAAGATTAATAC 2230
Db 1221 GAACCTCTGCTCTTGCACCCCTGGCGCAGTTCAGAGTGTCAACCCGGAAGGCTGA--- 1277
QY 2231 AGACGAGAAACATACCATTGAGCAACCTGCTCTTTTCAGTTTCAAAACAGTCACCAAC 2290
Db 1278 -----CAAAGATCACCAAC 1292
QY 2291 AATATCAACATCTAAATGGTTTACCCAAATCTATTGTAAGACACCAAGCAGCAATAC 2350
Db 1293 AATATCAGTTCTTAAATGGCTTGATCCAAAGTCTGCTTGTGAGACACCTAGTAGCAGCTC 1352
QY 2351 CTTGGATGATTAATGAGTGTGTTTGAACCTCAGTTGTAAGAAATGACTTTCCACTGC 2410
Db 1353 CTTGGATGATTAATGAAATGTTTGAAGACTCCAGTTGTAAGAAATGACTTTCCACTGC 1412
QY 2411 TTGTCAAGTTGTCAACACCTTATGCCCAACCTGCTGTTTCAGCAGCAACAGCATCAAAAT 2470
Db 1413 CTGTCCATCATACACCTTACAGCCAGCTTGCCCGCTCCAGCAGCAACAGCAGAGG 1472
QY 2471 ACTTGCCACTCCATCTCAAAATTTACAGGTTTACGTTTACGTTTACGTTTACGTTTACGTT 2530
Db 1473 ACTCAGCACTCCTCTTCAAGCTTGCAGATTTTCAGGTTCTTTCATCAATAAATGAATGCAT 1532
QY 2531 TTCGGTTAAAGGAAGATTTTATCCATATTAAGCAGATGAGGAGTTCAGGTTTCAAGCAA 2590
Db 1533 TTCAGTTAACGGAAGATTTTATCCATATTAAGCAGATGAGGAGTTCAGGTTTCCAGTAA 1592
QY 2591 GGTATTTTCAAGTGTGTTAAATGAAGAAACAGATATGCTATATAAATATGTAACCTTAGA 2650
Db 1593 GGTGTTTCAGGTTTGAATGAGAAACAGATAAAGCTATCAATATGTAACCTTAGA 1652
QY 2651 AGACGAGATTAACCAACTCTTGTAGTTTACCGGAAGAAATGACTTATTTGTAATAACT 2710
Db 1653 AGACGCGGATAGCCAAACTATTGAGAGCTACCGCAACAGATAGCGTTTGTGAACAACT 1712

QY 2711 ACAACAAACACAGTGTATGATCATCCGACTTTATGATTAATGAATCAACGACCACTACAT 2770
Db 1713 ACAGCAACACAGTGTATGATCATCCGCTCTATGATTAATGAATCAACGACCACTACAT 1772
QY 2771 CTACATGTAATGAGTGTGGAATTAATTCATCTTAAATAGTTGGCTTAAAGAGAAAAATC 2830
Db 1773 CTACATGTAATGGAATGTGGAACATTCAGCTTAATAGTTGGCTTAAAGAGAAAAATC 1832
QY 2831 CATTCATCCATGGGAACGACAGTGTACTGGAATAATGTTAGAGGAGTTCACACAT 2890
Db 1833 CATCAATCCATGGGAACGACAGTGTACTGGAATAATGTTAGAGGAGTTCACACAT 1892
QY 2891 CATCAACATGTCATGTTTCAAGTGTACTTAAACACAGCTAACTTTCTGATAGTTGATGG 2950
Db 1893 CCATCAGCATGTTATTTTCATAGTGAATCTGAAGCTCTGCTAACTTTGATAGTGAATGG 1952
QY 2951 AATGCTAAGCTAATTAATTTTGGGATTTGGAACCAATATGMAACAGATACACAGTGT 3010
Db 1953 AATGCTAAGCTAATTAATTTTGGGATTTGGAACCAATATGMAACAGATACACAGTGT 2012
QY 3011 TGTAAAGATTTCTCAGGTTGGCACAGTTAAATTAATGTCACCAAGCAATCAAAAGATAT 3070
Db 2013 TGTAAAGATTTCTCAGGTTGGCACAGTTAACTATATGSCCCAGAGCAATCAGAGATAT 2072
QY 3071 GTCTTCCTCCAGAGAAATGGAATCTAAGTCAAGATAAGCCCAAAAGTGTGTTG 3130
Db 2073 GTCTTCCTCCAGAGAAATTTGGAATCAGGACCAAGTAACTCCAGAGTGTGTTG 2132
QY 3131 GTCCCTAGAGTGTATTTGCTATATGATTTACGGGAACCAACCATTTACAGCATATAT 3190
Db 2133 GTCCCTAGAGTGTATTTGCTATATGATTTACGGGAACCAACCATTTACAGCATATAT 2192
QY 3191 TAATCAGATTTCTAAATTAATGATGCTCAATTAATGATCAATTAATTAATTAATTAAT 3250
Db 2193 CAATCAGGTTCTTAAACTGCAGCCATTAATCAACCTGCTCATGATGATTAATTTCCCGA 2252
QY 3251 TAATCAGAGAAAGATCTTCAAGATGTGTAAAGTGTGTTTAAAGAGGACCAAAACA 3310
Db 2253 GATTTCCGAAAAAGATCTTCGAGAGCTGTAAAGTGTGTTTAAAGAGGACCAAAACA 2312
QY 3311 GAGGATATCCATCTCTGAGCTCTGCTCATCATGTTCAAAATTAATTAATTAATTAAT 3370
Db 2313 GAGGATATCTATCCCTGAGCTTCTCACATATCCGTTATGTTCAATTCAGCCCATCCAG 2372
QY 3371 TAACCAATGCGCAAGGGAACCACTGGAAGAAATGAATATGTTCTGGGCAACTGTTGG 3430
Db 2373 CAGCCAAATGCGTAGGGAGCCACTGATGAATGAATATGTTGTTGGTCAACTGTTGG 2432
QY 3431 TCTGATTTCTCTTAACTCCATTTTCAAGCTCTCTAAA 3467
Db 2433 TCTGATTTCTCTTAACTCCATTTTGAAGAACTGCAAAA 2469

RESULT 2

AK049200 1841 bp mRNA linear HTC 20-SEP-2003
LOCUS Mus musculus ES cells cDNA, RIKEN full-length enriched library
DEFINITION clone: C330012K12 product: DUAL SPECIFICITY PROTEIN KINASE ITK (EC 2.7.1.1-) (BSK) (PYT) homolog [Mus musculus], full insert sequence.
ACCESSION AK049200
VERSION AK049200.1 GI:26339883
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

```

2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
AUTHORS Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes
MEDLINE Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 20499374
1 11042159
3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
AUTHORS Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
JOURNAL sequencing pipeline with 384 multipillar sequencer
MEDLINE Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 20530913
4 11076861
4 The RIKEN Genome Exploration Research Group Phase II Team and the
AUTHORS FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
AUTHORS Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
JOURNAL of 60,770 full-length cDNAs
REFERENCE Nature 420, 563-573 (2002)
6 (bases 1 to 1841)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,P., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Saeki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akaira,S.,
Takeda,Y., Tanaka,T., Tonaru,A., Toyota,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
source Location/Qualifiers
1..1841
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:C330012K12"
/db_xref="MGI:2416445"
/db_xref="taxon:10090"
/cclone="C330012K12"
/cell_type="ES cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
85..51839
CDS

```

```

/note="unnamed protein product; DUAL SPECIFICITY PROTEIN
KINASE TTK (EC 2.7.1.-) (ESK) (PYT) homolog [Mus musculus]
(SWISSPROT|P35761, evidence: FASTY, 99.9%ID, 96.7%length,
match=2409)
putative"
/codon_start=1
/protein_id="BAC33605.1"
/db_xref="GI:26339884"
/translation="MEAEELIGSSVTIDISMSKWRDINKINEDCTDELSKICADH
TETVQMRVGNLTENFWLFLKLEKNSFPLNDLLKLGYSQAIEVLPPKYGQN
ESFARIQVRLAEKAIQEPDDARDYQVARENCKKFAFVHVSQAQFELSGNKKSSQ
LLHKVATGAVPLQMLNHLKQQLLPREDKKSASTVLSQAQFVSSLGNV
QNRISCESGQGAARLVXENLPPODAVRHONPFKQTHAAKRSQPFKQFVNLIN
SPDYKVTDSANTQLTRKSGSPDRAILPGSPRSGSDSYELRGKPIOTIYLKQSL
VSNKSELMDLIALSKSDISLTKLEETKPEIARRRPMQWQSTRKPECVFQNPAPF
APLRHVPDVTPKADKSPPIVSFKMLPKSACETPSSSLDDYMKKPKTPVVRNDPFP
ACPSSTYSLARLQOQQQLSTPLQSLQISGSSINECISVNGRIYSILKQIGGG
SSKVFQVLNKKQINAIKYNLEBDADSTQIESYRNEIAFLNKLQHSKILRLDYVEI
TEQVIYVMVECCNIDLSNW"
Query Match 27.1%; Score 1047.2; DB 11; Length 1841;
Best Local Similarity 76.5%; Pred. No. 3.1e-211;
Matches 1410; Conservative 0; Mismatches 353; Indels 81; Gaps 7;
QY 971 TGTAGAAATGGAATCCGAGGATTTAAAGTGGCAGAGAAATGCAATTCATTGATTCCTAATGAA 1030
DB 78 TTTAGAAATGGAGCTGAGAGTTAATGGCAGCAGTGTGACGATTCATTGATTCCTAATGAG 137
QY 1031 CAAAGTGGAGACATTTAAAATAAGTTTAAAATAAGTCTTACTGATGAACCTTACTGATGAACCTT 1090
DB 138 CAAAATGAGAGATATTAAAAATAAG---ATAAATGAAAGACTGTACTGATGAGCTAAGCTT 194
QY 1091 GAATAAAATTTCTGCTGATCTACAGATAACTCGGGAAGTGTAAACCAAAATTTATGATGAT 1150
DB 195 GTCTAAAATCTGTG-----CCGATCACAGGAAACTGTTAACCAATTTATGAGGCT 245
QY 1151 GCGAAACACCCAGAGAGACTGGTTGAGTTTGTGCTCAAACTAGAGAAAAACAGTGTTC 1210
DB 246 TGGGAACACCCAGAGAACTGGTTGAAATTTTGTGCTGAAACTAGAGAAAAACAGCTCAAC 305
QY 1211 CTTAAGTATGCTCTTTTAAATAAATGATGCTGTCTAGCTCAAGCAATTCAGAGCTT 1270
DB 306 TCTAAATGAACATCTTTTAAATAAGCTGATGTTGGTGGTATAGTCAAGCGATTGAAGTACT 365
QY 1271 TCCCCAGATAATATGCGCAAAATGAGAGTTTGTGCTAGAAATTCAGTGAAGATTGCTGA 1330
DB 366 TCCTCCAGATAAATACGCCAGATGAGAGCTTTGCTCGAATACAAGTGAAGACTTGCTGA 425
QY 1331 ATTTAAAGCTATTCAAGGCCAGATGATGACGTGACTTCTTTCAATGCGCAGAGACAAA 1390
DB 426 ACTAAAAGCTATTCAAGAGCGCTGATGATGCGCGTACTTCCAGATGCGCAGGGA 485
QY 1391 CTGCAAGAAATTTCTTTTGTTCATATATCTTTTGCACAAATTTGCACTGCACAAAGTAA 1450
DB 486 CTGCAAGAGTTTCTTTTGTGACAGTATCTTTTGCACAGTTTGAATCTGCTCAAGCAA 545
QY 1451 TGTCAAAAAAGTAAACAACTTTCTTCAAAAGCTGTAGACGCTGAGAGAGTACCACTAGA 1510
DB 546 TCTTAAAAAAGTGAAGCAGCTTCTTCAAAAGCTGTAGAGACTGGGCGGTGCGCGTGA 605
QY 1511 AATCTGGAAATTCCTCGGGGATTTAAACCTCCAAAAAAGCAGCTGCTTTTCAGAGGA 1570
DB 606 GATGCTGGAGCGGCCATGCGTAACTTACCTCCAGAAAAGCAGCTGCTTCGGAGGA 665
QY 1571 GGAAGAAGAAATTTATCAGCATCTACCGTATTAACTGCCCAAGAAATCATTTCCCGTTC 1630
DB 666 GGACAAGAAGAGTGTGTGAGCATCGACAGTACTAAGTGCCCAAGAGCGGTCTTCCAGCTC 725
QY 1631 ACTTGGGCTTTTACAGAAATGAGACACAGTGTGTTGATTCAGAGGAGCAGACTACTAAAGC 1690
DB 726 ACTTGGAAATGTACAGAATCGAGCATCAGCTGTGAGTCCAGAGGACAGCTGGGCGAGC 785

```

QY 1691 CAGGTTTTTATATGAGAGAAATGCGCCACCAAGATGCGAAATAGGTTACCGGAATTC 1750
Db 786 CAGGGTTTTATATGAGAGAAATGCGCTCCACCAAGATGCGGAAGTGAAGGATCAAAACCC 845
QY 1751 ATTGAGACAACTACAAAATTAACAGTCTATGCCATTTGGAGAGATCCAGTTAACT 1810
Db 846 CTTCAAGCAGACTCAGCAGCTAAACCGTCTATGCCCTTTGGAAAATGCCAGTCAATCT 905
QY 1811 TCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTGATCTGTTTATGAA 1870
Db 906 TCTAAACAGCCAGATTCTATGTGAAGACAGATGATTCAGTTGTGATCTGTTTATGAA 1870
QY 1871 AAGACAACTCTATAGATCAGAAATCGCGAGATTTGGTTGCTGCTGATCTAAACCAAGTGG 1930
Db 956 AAGAAAGGGTCCAGACACAGAA---CCGAGACGCGATTCGCGCGCTCCAGACCACTGG 1022
QY 1931 AAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCAATTTCAAGGAACC 1990
Db 1023 CAGTGATTCCTATGACTGAGAGGTTAAAGCCATTCAAACTATCTATTTGAAGACTC 1082
QY 1991 TCTGGTGCAGATGAAGAGATTCGAACTTATTTATCTGATTCATTAACCCCTGAAGAA 2050
Db 1083 TTTGGTGTCCAATGAAAAGATTTCTGAAC---TTATGTCTGATTTAATAGCTTTGAAGAG 1139
QY 2051 TAAACGGAATCAAGTCTCTAGTCTAAATTAAGAGAACTAAAGAGTATCAAGAACGAGA 2110
Db 1140 TAAACAGATTCAG---TCTAACAATTTGGAAGAACTA-----AGCCAGA 1184
QY 2111 GGTTCAGAGAGTACACGAAACAGTGGCAAGCTTAAGAGAAAGTCAGAGTGTATTAACCA 2170
Db 1185 GATTGCAGAAAGAGCCCATGTCAGTGGCAGTCTACCAAGAAAGCCGAGTGTGTGCCA 1244
QY 2171 GAATCTCTGCTGATCTTCAATATCACTGGCAGATTCGGAGTACCGCGAAGAAAGTATAC 2230
Db 1245 GAACCTCTGCTGCTTTCACCCCTGCGGCAGCTTCCAGATGTACCCCGAGGCTGA--- 1301
QY 2231 AGAGCAGAAACATACCACTTTTGAAGCACTGTCTTTTCAAGTTTCAAAAAGTCACCAACC 2290
Db 1302 -----CAAGAGTACCAACC 1316
QY 2291 AATATCAACATCTAAATGGTTTGACCCAAATCTATTGTGAAGACACCAAGCAGCAATAC 2350
Db 1317 AATATCAGTTCCTAAATGGCTGTGATCAAAAGTCTGCTGTGAGACACCTAGTAGCAGCTC 1376
QY 2351 CTTGGATGATTCAGTACGCTGTTTGTAGAACTCCAGTGTGAAGAAATGACATTTCCACCTGC 2410
Db 1377 CTTGGATGATTCAGTAAATGTTTATAGACTTCCAGTGTGAAGAAATGACATTTCCACCTGC 1436
QY 2411 TTGTCAGTGTCAACACCTTATGCGCAACCTGCTGTTTCCAGCAGCAACAGCAGATCAAT 2470
Db 1437 CTGTCCATCATCAACACCTTACAGCCAGCTTGCCGCGCTCCAGCAGCAACAGCAGCAGG 1496
QY 2471 ACTTGCCACTCCACTTCAAAATTTACAGGTTTATAGCATCTCTTTCAGCAATGATGCAT 2530
Db 1497 ACTCAGCACTCTCTTCAAGCTTGCAGATTTCCAGTTCCTTCATCAATAAATGAATGCAT 1556
QY 2531 TTCGGTTAAAGGAAGAAATTTATTCATATTAAGCAGATAGGAAGTGGAGTTTCAAGCAA 2590
Db 1557 TTCAGTTAACGGAAGAAATTTATTCATATTAAGCAGATAGGCAATGAGGTTCCAGTAA 1616
QY 2591 GGTATTTCAAGTGTAAATGAAGAAAGAAACAGATATGCTATATAAAATATGTGAACCTAGA 2650
Db 1617 GGTGTTTCAGGTTATGATGAGAAAAAAGACATAAACGCTATCAATATGTGAACCTAGA 1676
QY 2651 AGAGCAGATACCAACCTCTTGTAGTTTACCGGAACGAAATAGCTTATTTGAATAACT 2710
Db 1677 AGACCCGATAGCCAACTATGAGAGCTACCGCAACGAGATAGCGTTTGTGAACAACT 1736
QY 2711 ACAACAAACAGTGAATAGATCATCCGACTTTTATGATTAATGAATCAACGCAAGTACAT 2770
Db 1737 ACAGCAACAGTGAATAGATCATCCGCTCTATGATTAATGAATCAACGCAAGTACAT 1796
QY 2771 CTACATGTAATGGAGTGTGGAATATTTGATCTTAATAGTTGGC 2814

Db 1797 CTACATGTAATGGAATGTGGAACATTTGACCTAAATAGTTGGC 1840

RESULT 3
EX415805
LOCUS
DEFINITION
EX415805 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YE07
5-PRIME, mRNA sequence.
EX415805
VERSION
EX415805.1 GI:30634668
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Li W.B., Gruber C., Jessee J. and Polayes D.
1 (bases 1 to 1002)
Full-length cDNA libraries and normalization
Unpublished (2001)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3657.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0CAP008AC04Q1&cluster=3657.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0CAP008AC04Q1.

FEATURES
Location/Qualifiers
1..1002
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP008YE07"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 23.4%; Score 906.2; DB 13; Length 1002;
Best Local Similarity 98.8%; Pred. No. 2.5e-181;
Matches 916; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 1999 CAGATCAAGAGAGTCTGAACTTATTACTGATTCAATACCTCGAGATAAAACGG 2058
Db 70 CGGATGAAAGAGTCTGAACTTATTACTGATTCAATACCTCGAGATAAAACGG 129
QY 2059 AATCAAGTCTTCTAGCTAAATAGAGAACTAAAGAGATATCAAGAACCCAGAGTTCAG 2118
Db 130 AATCAAGTCTTCTAGCTAAATAGAGAACTAAAGAGATATCAAGAACCCAGAGTTCAG 189
QY 2119 AGAGTACACAGAAACAGTGGCAGCTAAGAGAACTCAGAGTGTATTACCAAGATCTCTG 2178
Db 190 AGAGTACACAGAAACAGTGGCAGCTAAGAGAACTCAGAGTGTATTACCAAGATCTCTG 249
QY 2179 CTGCATCTTCAATCACTGGCAGATTCGCGAGTTAGCCGAAAGTTAATACAGAGCAGA 2238
Db 250 CTGCATCTTCAATCACTGGCAGATTCGCGAGTTAGCCGAAAGTTAATACAGAGCAGA 309
QY 2239 AACATACCACTTTTGAAGCAACCTGCTTTTTCAGTTTCAAAAAGTCAACCAATATCAA 2298
Db 310 AACATACCACTTTTGAAGCAACCTGCTTTTTCAGTTTCAAAAAGTCAACCAATATCAA 369
QY 2299 CATCTAAATGGTTTGAACCCCAAAATCTATTGTAAGACACCAAGCAGCAATACCTGGATG 2358

```

Db 370 CATCTAAATGGTTGACCCAAATCTATTTGTAAAGACCAAGCAGCAATACCTTGGATG 429
QY 2359 ATTACATGAGCTGTTTAGAAGTCTCAGTTGTAAAGAAATGACTTTCCACCTGCTTGTTCAGT 2418
Db 430 ATTACATGAGCTGTTTAGAAGTCTCAGTTGTAAAGAAATGACTTTCCACCTGCTTGTTCAGT 489
QY 2419 TGTCAACACCTTATGGCAACCTGCTTGTTCAGCAGCAACAGCATCAAAATCTTGCCA 2478
Db 490 TGTCAACACCTTATGGCAACCTGCTTGTTCAGCAGCAACAGCATCAAAATCTTGCCA 549
QY 2479 CTCACATCTCAAAATTTACAGTCTTTCAGCAATCTTTCAGCAATAGATGATTCGCTTA 2538
Db 550 CTCACATCTCAAAATTTACAGTCTTTCAGCAATCTTTCAGCAATAGATGATTCGCTTA 609
QY 2539 AAGGAAGAATTTATTCATATTAAGCAGATAGGAAGTGGAGTTCAAGCAAGGTATTTTC 2598
Db 610 AAGGAAGAATTTATTCATATTAAGCAGATAGGAAGTGGAGTTCAAGCAAGGTATTTTC 669
QY 2599 AGGTGTTAAATGAAAGAAACAGATATATCTATATAATATGTGAATTAAGTAAAGTAAAGCAG 2658
Db 670 AGGTGTTAAATGAAAGAAACAGATATATCTATATAATATGTGAATTAAGTAAAGTAAAGCAG 729
QY 2659 ATAAACCAAACTCTTGATAGTTTACCGGAACGAAATAGCTTATTTGAATTAAGTAAAGTAAAGCAG 2718
Db 730 ATAAACCAAACTCTTGATAGTTTACCGGAACGAAATAGCTTATTTGAATTAAGTAAAGTAAAGCAG 789
QY 2719 ACAGTGATAGATCATCGGACTTTATGATTAAGTAAATCAGGACCGAGTACATCTACATGG 2778
Db 790 ACAGTGATAGATCATCGGACTTTATGATTAAGTAAATCAGGACCGAGTACATCTACATGG 849
QY 2779 TAATGGAGTGGGAAATATTGATCTTAATAGTTCGCTTAAAGGAAAGAAATATCCATTGATC 2838
Db 850 TAATGGAGTGGGAAATATTGATCTTAATAGTTCGCTTAAAGGAAAGAAATATCCATTGATC 909
QY 2839 CATGGGAACGCAAGAGTTACTGGAAAATATGTTTAGAGGAGTTTCAACAATCCATCAAC 2898
Db 910 CATGGGAACGCAAGAGTTACTGGAAAATATGTTTAGAGGAGTTTCAACAATCCATCAAC 968
QY 2899 ATGGCATTTTTCACAGTCACTTAAAC 2925
Db 969 ATGGCATTTTTCACAGTCACTTAAAC 995

RESULT 4
LOCUS AUI32784
DEFINITION AUI32784 NT2RP4 Homo sapiens cDNA clone NT2RP4000532 5', mRNA
sequence.
ACCESSION AUI32784
VERSION AUI32784.1 GI:10993323
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 887)
Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J.,
Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuho,Y. and
Isogai,T.
HRI human cDNA project (Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y.,
Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T.,
Sugano,S., Masuho,Y., Isogai,T.)
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
152-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and

```

```

FEATURES
source
    Helix Research Institute.
    Location/Qualifiers
        1..887
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="NT2RP4000532"
        /cell_type="teratocarcinoma"
        /cell_line="NT2"
        /clone_lib="NT2RP4"
        /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
        cells after 2-weeks retinoic acid (RA) induction"

ORIGIN
Query Match      22.0%; Score 951.8; DB 9; Length 887;
Best Local Similarity 98.8%; Pred. No. 8.5e-170;
Matches 878; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 1105 CTGATACACAGATAAATCGGGAAGTGTAAACCAATTTATGATGATGCAACCAACCCAG 1164
Db 1 CTGATACACAGATAAATCGGGAAGTGTAAACCAATTTATGATGATGCAACCAACCCAG 60
QY 1165 AGGACTGGTTGAGTTTGTTCCTCAAACTAGAGAAAACAGTGTTCGCTAAGTGATGCTC 1224
Db 61 AGGACTGGTTGAGTTTGTTCCTCAAACTAGAGAAAACAGTGTTCGCTAAGTGATGCTC 120
QY 1225 TTTTAAATTAATGATGGTTCGTACAGTCAAGCAATTTGANGCGCTTCCCCAGATAAAT 1284
Db 121 TTTTAAATTAATGATGGTTCGTACAGTCAAGCAATTTGANGCGCTTCCCCAGATAAAT 180
QY 1285 ATGGCAAAATCAGAGTTTTCGTAGAATTTCAAGTCAGATTTTCGTAATTTAAACCTATT 1344
Db 181 ATGGCAAAATCAGAGTTTTCGTAGAATTTCAAGTCAGATTTTCGTAATTTAAACCTATT 240
QY 1345 AAGAGCCAGATGATGACGCTGACTACTTTTCAATTTGGCCAGAGCAAACTGCAGAAATTTG 1404
Db 241 AAGAGCCAGATGATGACGCTGACTACTTTTCAATTTGGCCAGAGCAAACTGCAGAAATTTG 300
QY 1405 CTTTGTGTTTCATATATCTTTTCACAAATTTGAACGTGTCACAGGTAATGTCAAAAAAGTA 1464
Db 301 CTTTGTGTTTCATATATCTTTTCACAAATTTGAACGTGTCACAGGTAATGTCAAAAAAGTA 360
QY 1465 AACAACTTTCTTCAAAAAGCTGTAGAAGCTGGAGCAGTACCACTAGAAATGCTGGAATTTG 1524
Db 361 AACAACTTTCTTCAAAAAGCTGTAGAAGCTGGAGCAGTACCACTAGAAATGCTGGAATTTG 420
QY 1525 CCCTGCGGAATTTAAACCTCCAAAAGAGCAGCTGCTTTCAGAGGAGGAAGAAAGAAATTT 1584
Db 421 CCCTGCGGAATTTAAACCTCCAAAAGAGCAGCTGCTTTCAGAGGAGGAAGAAAGAAATTT 480
QY 1585 TATCAGCATCTACGGTATTAACCTGCCAAGAAATCATTTTCGGTTCACCTTGGGCAATTTC 1644
Db 481 TATCAGCATCTACGGTATTAACCTGCCAAGAAATCATTTTCGGTTCACCTTGGGCAATTTC 540
QY 1645 AGAATAGAGAACACAGTTTGTGATTTCCAGAGCAGACAGTACTTAAAGCCAGGTTTTATATG 1704
Db 541 AGAATAGAGAACACAGTTTGTGATTTCCAGAGCAGACAGTACTTAAAGCCAGGTTTTATATG 600
QY 1705 GAGAGAAATGCTCCACCAAGATGCAAAATAGGTTTACCGGAATTCATTTCAGACAAAATA 1764
Db 601 GAGAGAAATGCTCCACCAAGATGCAAAATAGGTTTACCGGAATTCATTTCAGACAAAATA 660
QY 1765 ACAAACTAAACAGTCAATGCCCATTTTGGAGAGTCCCGAGTTTAACTTCTTAAATAGCCAG 1824
Db 661 ACAAACTAAACAGTCAATGCCCATTTTGGAGAGTCCCGAGTTTAACTTCTTAAATAGCCAG 720
QY 1825 ATTGTGATGTGAAGACAGATGATTCAGTTGTACCTGTTTATGAAAGACAAACCTCTA 1884
Db 721 ATTGTGATGTGAAGACAGATGATTCAGTTGTACCTGTTTATGAAAGACAAACCTCTA 779
QY 1885 GATCAGAAATGCCAGATTTGTTGCTGCTGATCTAAACCAAGTGGAAATGATTCCTGTG 1944
Db 780 GATCAGAAATGCCAGATTTGTTGCTGCTGATCTAAACCAAGTGGAAATGATTCCTGTG 839

```

```

QY 1945 AATTAGAAATTTAAAGTCTGTTCAAAATAGTCAATTTCAAGGAACCTCT 1993
Db 840 AATTAGAAA-TTAAAGTCTGGTCAAAATAGTCAATTTCAAGGGAACCT 897

RESULT 5
LOCUS BQ420892 866 bp mRNA linear EST 23-MAY-2002
DEFINITION AGENCOURT_7802344 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6041805
5', mRNA sequence.
ACCESSION BQ420892
VERSION BQ420892.1 GI:21116207
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM3280 row: c column: 22
High quality sequence stop: 580.
FEATURES
Location/Qualifiers
1..866
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6041805"
/tissue types="embryonal carcinoma, cell line"
/lab host="DH10B (phage-resistant)"
/clone lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 21.1%; Score 816.4; DB 13; Length 866;
Best Local Similarity 98.0%; Pred. No. 2.7e-162;
Matches 826; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 967 TTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAATTGACAAATGATTCCTATAA 1026
Db 1 TTTTCTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAATTGACAAATGATTCCTATAA 60

QY 1027 TGAACAAAGTGAGAGACATTTAAATTAAGTTTAAATTAAGAGACCTTACTGATGAACATA 1086
Db 61 TGAACAAAGTGAGAGACATTTAAATTAAGTTTAAATTAAGAGACCTTACTGATGAACATA 120

QY 1087 GCTTGAATTAATTTCTGCTGATCTACTACAGATACTCGGGAACGTGTTAAACCAATTAAGA 1146
Db 121 GCTTGAATTAATTTCTGCTGATCTACTACAGATACTCGGGAACGTGTTAAACCAATTAAGA 180

QY 1147 TGATGGCAAAACCAACCCAGAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAAACAGTG 1206
Db 181 TGATGGCAAAACCAACCCAGAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAAACAGTG 240

QY 1207 TTCGCTAAGTGATGCTCTTTTAAATTAATTAAGTTGCTGTTACAGTCACGCAATTCAG 1266
Db 241 TTCGCTAAGTGATGCTCTTTTAAATTAATTAAGTTGCTGTTACAGTCACGCAATTCAG 300

1267 GCCTTCCCCAGATAAAATATGCGCCAAAATGAGAGTTTCTAGATTCAAGTGAGATTG 1326
301 GCCTTCCCCAGATAAAATATGCGCCAAAATGAGAGTTTCTAGATTCAAGTGAGATTG 360
1327 CTGAATTAAGAGCTATTCAGAGCCAGATGATGACGTCGACTACTCTTTCAAAATGGCCAG 1386
361 CTGAATTAAGAGCTATTCAGAGCCAGATGATGACGTCGACTACTCTTTCAAAATGGCCAG 420
1387 CAAACTGCAAGAAATTTCTCTTTTGTTCATATATCTTTTGGCACAATTTGCAACTGTGACAAG 1446
421 CAAACTGCAAGAAATTTCTCTTTTGTTCATATATCTTTTGGCACAATTTGCAACTGTGACAAG 480
1447 GTATATGTCAAAAAAGTAACAACATCTTCTCAAAAAGCTGTAGAACGTCGAGCAGTACCAC 1506
481 GTATATGTCAAAAAAGTAACAACATCTTCTCAAAAAGCTGTAGAACGTCGAGCAGTACCAC 540
1507 TAGAATATGCTGGAATTTCCCTGCGGAAATTTAAACCTCCAAAAGAGCAGTGCTTTCCAG 1566
541 TAGAATATGCTGGAATTTCCCTGCGGAAATTTAAACCTCCAAAAGAGCAGTGCTTTCCAG 600
1567 AGGAGAGAAAAGAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAATCAITTTCCG 1626
601 AGGAGAGAAAAGAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAATCAITTTCCG 660
1627 GTTCACTTGGGCAATTTACAGAAATAGGAACAACAGTTGTGATTCAGAGAGCAGACTACTA 1686
661 GTTCACTTGGGCAATTTACAGAAATAGGAACAACAGTTGTGATTCAGAGAGCAGACTACTA 720
1687 AAGCCAGAGTTTATATGAGAGAGAACATGCGCACCAAGATGCGAAGATCGAAGATAGTTACCGGA 1746
721 AAGCCAGAGTTTATATGAGAGAGAACATGCGCACCAAGATGCGAAGATCGAAGATAGTTACCGGA 780
1747 ATTCAATGAGACAAACTAAACAACTAAACAGTCACTGCGCAATTCGAGAGAGTCCCAAGTTA 1806
781 ATTCAATGAGACAAACTAAACAACTAAACAGTCACTGCGCAATTCGAGAGAGTCCCAAGTT 840
1807 ACC 1809
841 AAC 843

RESULT 6
LOCUS AU124386 916 bp mRNA linear EST 01-AUG-2002
DEFINITION AU124386 NT2RM2 Homo sapiens cDNA clone NT2RM2002111 5', mRNA
sequence.
ACCESSION AU124386
VERSION AU124386.1 GI:10949102
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 916)
AUTHORS Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,
Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and
Ishigai,T.
TITLE HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,
Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T.,
Suzuki,Y., Sugano,S., Ishigai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
Location/Qualifiers

```

```
source
1. .916
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RM2002111"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_lib="NT2RM2"
/note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
precursor cells"

ORIGIN
Query Match 21.0%; Score 812.8; DB 9; Length 916;
Best Local Similarity 95.7%; Pred. No. 1.5e-161;
Matches 877; Conservative 0; Mismatches 31; Indels 8; Gaps 4;

QY 1649 TAGGAACAACAGTTGTGATTCAGAGGACAGACTACTAAAGCCAGGTTTATATGAGGA 1708
Db 1 TAGGAACAACAGTTGTGATTCAGAGGACAGACTACTAAAGCCAGGTTTATATGAGGA 60
QY 1709 GAACATGCCACCAAGATGAGGAAATAGTTACCGGAATTCATGAGCAAACTACAA 1768
Db 61 GAACATGCCACCAAGATGAGGAAATAGTTACCGGAATTCATGAGCAAACTACAA 120
QY 1769 AACTAAACAGTCATGCCATTTGGAAGAGTCCAGTTAACTTTCTAAATAGCCAGATTG 1828
Db 121 AACTAAACAGTCATGCCATTTGGAAGAGTCCAGTTAACTTTCTAAATAGCCAGATTG 180
QY 1829 TGATGTGAACAGATGATTCAGTTGACCTGTTTATGAAAGACAACTCTAGATC 1888
Db 181 TGATGTGAACAGATGATTCAGTTGACCTGTTTATGAAAGACAACTCTAGATC 240
QY 1889 AGAATGCCAGATTTGTTGTCCTGATCAACCAAGTGGAAATGATTCCTGTGAATT 1948
Db 241 AGAATGCCAGATTTGTTGTCCTGATCAACCAAGTGGAAATGATTCCTGTGAATT 300
QY 1949 AAGAAATTAAGTCTGTTCAAAATAGTCAATTCAGGAACCTCTGGTGCAGATGAAA 2008
Db 301 AAGAAATTAAGTCTGTTCAAAATAGTCAATTCAGGAACCTCTGGTGCAGATGAAA 360
QY 2009 GAGTTCTGAATCTATTATTAATGATCAATCAACCTGAAGATTAACCGAATCAAGTCT 2068
Db 361 GAGTTCTGAATCTATTATTAATGATCAATCAACCTGAAGATTAACCGAATCAAGTCT 420
QY 2069 TCTAGCTAAATAGAGAACTAAAGAGTATCAAGAACAGAGTTCCAGAGAGTAACCA 2128
Db 421 TCTAGCTAAATAGAGAACTAAAGAGTATCAAGAACAGAGTTCCAGAGAGTAACCA 480
QY 2129 GAAACAGTGCAGCTAAGAGAAAGTCAAGTGTATTAACCAATCCTGCTGATCTTC 2188
Db 481 GAAACAGTGCAGCTAAGAGAAAGTCAAGTGTATTAACCAATCCTGCTGATCTTC 540
QY 2189 AAATCACTGGCAGATCCGGAGTTAGCCGAAAGTTAATACAGAGCAGAAATACCA 2248
Db 541 AAATCACTGGCAGATCCGGAGTTAGCCGAAAGTTAATACAGAGCAGAAATACCA 600
QY 2249 TTTTGACCAACCTGCTTTTCAAGTCAACAGTCAACCAATATCAATCAATCAATG 2308
Db 601 TTTTGACCAACCTGCTTTTCAAGTCAACAGTCAACCAATATCAATCAATCAATG 660
QY 2309 GTTTGACCCAAATCTATTGTTAAGACACCAACAGCAATACCTTGGATGATTAATGAG 2368
Db 661 GTTTGACCCAAATCTATTGTTAAGACACCAACAGCAATACCTTGGATGATTAATGAG 720
QY 2369 CTGTTTATAGAACTCCAGTTGT-AAAGATGACTTTTCCACTGCTGTTGCTGAGTCAAC 2427
Db 721 CTGTTTATAGAACTCCAGTTGTAAAGATGACTTTTNCACCTGCTG-CAGTTGTCAAC 779
QY 2428 CTTATGCCAACCTGCTGTTTCCAGCAGCAACAGCATCAATATCTTGGCACT--CCACT 2485
Db 780 CTTATGCCAACCTGCTGTTTCCAGCAGCAACAGCATCAATATCTTGGCACTTCCACT 839
QY 2486 TCAAAATTTACAGGTTTATGATCTCTTCAGCAAAA-----TGAATGCATTTCCGTTAAG 2541
```

```
Db 840 CAAAATTTACAGTTTAACTCTCTTAAACCAATGGAATGCCCTTCCTTCNGGTTAAG 899
QY 2542 GAAGAAATTTATTCAT 2557
Db 900 GACGAATTTTCTCT 915

RESULT 7
BQ929926 931 bp mRNA linear EST 20-AUG-2002
LOCUS AGENCOURT 8950097 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6471444
DEFINITION 5', mRNA sequence.
ACCESSION BQ929926
VERSION BQ929926.1 GI:22344957
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
MAMMALIA; Eukarya; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 931)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIA014003 Row: m column: 13
High quality sequence stop: 656.
FEATURES
Location/Qualifiers
1..931
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6471444"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb. "
```

ORIGIN

```
Query Match 21.0%; Score 811.4; DB 13; Length 931;
Best Local Similarity 94.7%; Pred. No. 3e-161;
Matches 869; Conservative 0; Mismatches 40; Indels 9; Gaps 3;

QY 2640 GTGAACCTTAGAAGACAGATTAACCAACTCTTGTATAGTTACCGAACGAAATAGCTTAT 2699
Db 1 GTGAACCTTAGAAGACAGATTAACCAACTCTTGTATAGTTACCGAACGAAATAGCTTAT 60
QY 2700 TTGAATAAATPACAAACACAGTGATAAGATCATCCGATTTATGATTTATGAATCAACG 2759
Db 61 TTGAATAAATPACAAACACAGTGATAAGATCATCCGATTTATGATTTATGAATCAACG 120
QY 2760 GACCAGTACATCTACATGGTAATCGAGTGGAAATATTCATCTTAATAGTTCGCTTAA 2819
Db 121 GACCAGTACATCTACATGGTAATCGAGTGGAAATATTCATCTTAATAGTTCGCTTAA 180
QY 2820 AAGAAAAATCCATTTGATCCATGGGAACGCAAGAGTTACTCGAAAAATATGTTAGAGCA 2879
Db 181 AAGAAAAATCCATTTGATCCATGGGAACGCAAGAGTTACTCGAAAAATATGTTAGAGCA 240
QY 2880 GTTCACACAAATCCATCAACATGTCATTTTCAGTGCATCTTAAACACAGCTACTTCTG 2939
Db 241 GTTCACACAAATCCATCAACATGTCATTTTCAGTGCATCTTAAACACAGCTACTTCTG 300
```

QY 2940 ATAGTTGATGAAGTCTAAAGCTAAATGATTTTGGGATTCGAAACCAAAATGCCACACAGAT 2999
 DB |||||||
 QY 301 ATAGTTGATGAAGTCTAAAGCTAAATGATTTTGGGATTCGAAACCAAAATGCCACACAGAT 360
 DB |||||||
 QY 3000 ACAACAAGTGTGTAAAGATTTCTCAGGTTGGCAGATTAATTATATGCCACAGAGCA 3059
 DB |||||||
 QY 361 ACAACAAGTGTGTAAAGATTTCTCAGGTTGGCAGATTAATTATATGCCACAGAGCA 420
 DB |||||||
 QY 3060 ATCAAGATATGTTCTCTCCAGAGAGATGGGAAATCTAAGTCAAGATAGAGCCCAAA 3119
 DB |||||||
 QY 421 ATCAAGATATGTTCTCTCCAGAGAGATGGGAAATCTAAGTCAAGATAGAGCCCAAA 480
 DB |||||||
 QY 3120 AGTGATGTTGGTCTCTAGGATGATTTTGTACTATATGATTTACGGGAAACACCAATTT 3179
 DB |||||||
 QY 481 AGTGATGTTGGTCTCTAGGATGATTTTGTACTATATGATTTACGGGAAACACCAATTT 540
 DB |||||||
 QY 3180 CAGCAGATATTAATCAGATTTCTAAATATACATGCCATTAATGATCTATCATGAATTT 3239
 DB |||||||
 QY 541 CAGCAGATATTAATCAGATTTCTAAATATACATGCCATTAATGATCTATCATGAATTT 600
 DB |||||||
 QY 3240 GAATTTCCCGATATTCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAGG 3299
 DB |||||||
 QY 601 GAATTTCCCGATATTCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAGG 660
 DB |||||||
 QY 3300 GACCCAAACAGAGATATCATTTCTGAGCTCTGCTCATCATATGTTCAAAATCAA 3359
 DB |||||||
 QY 661 GACCCAAACAGAGATATCATTTCTGAGCTCTGCTCATCATATGTTCAAAATCAA 720
 DB |||||||
 QY 3360 ACTCATCCAGTAAACCAATGGCCAGAGGACCACTGAAGAAATGAATATGTTCTGGC 3419
 DB |||||||
 QY 721 ACTCATCCAGTAAACCAATGGCCAGAGGACCACTGAAGAAATGAATATGTTCTGGC 780
 DB |||||||
 QY 3420 CAATTTGTTGCTGTAATTTCTCAATCTTGAAGCTGCTAAACCTTTATATGA - 3478
 DB |||||||
 QY 781 CAATTTGTTGCTGTAATTTCTCAATCTTGAAGCTGCTAAACCTTTATATGA 840
 DB |||||||
 QY 3479 --ACACTATAGTGTGGTGAAGTCATAAT--CTTCATCTCC----AGACTTTTGA 3530
 DB |||||||
 QY 841 ACATATAGTGTGGTGAAGTCATAATTCCTCCATCCAGGAACTTTTGA 900
 DB |||||||
 QY 3531 AAAAAAGGGGAAAAAA 3548
 DB |||||||
 QY 901 AAAAAAGGGGAAAAAA 918
 DB |||||||

RESULT 8

BX361758/c
 LOCUS 1059 bp mRNA linear EST 05-MAY-2003
 DEFINITION BX361758 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
 Homo sapiens cDNA clone CS0DJ004YA24 3-PRIME, mRNA sequence.
 ACCESSION BX361758
 VERSION BX361758.1 GI:30376559
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1059)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqre@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 3657.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DJ004BA12NP1&cluster=3657.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DJ004BA12NP1.

FEATURES
source

Location/Qualifiers
 1..1059
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DJ004YA24"
 /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /cell_line="JURKAT"
 /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCWSF6 vector. Library was normalized."

ORIGIN

Query Match 20.3%; Score 783.8; DB 13; Length 1059;
 Best Local Similarity 93.9%; Pred. No. 2.1e-155;
 Matches 820; Conservative 24; Mismatches 24; Indels 5; Gaps 3;
 QY 2950 GAATGCTAAAGCTAAATGATTTTGGGATTCGAAACCAAAATGCCACAGATCAACAAGTG 3009
 DB |||||||
 QY 871 GAATGCTAAAGCTAAATGATTTTGGGATTCGAAACCAAAATGCCACAGATCAACAAGTG 813
 DB |||||||
 QY 3010 TTGTTAAAGATTTCTCAGGTTGGCAGATTAATTATATGCCACAGAGCAATCAAGATA 3069
 DB |||||||
 QY 812 TTGTTAAAGATTTCTCAGGTTGGCAGATTAATTATATGCCACAGAGCAATCAAGATA 753
 DB |||||||
 QY 3070 TGTCTTCCTCCAGAGAGATGGGAAATCTAAGTCAAGATAGAGTACCCCAAGATGATGTT 3129
 DB |||||||
 QY 752 TGTCTTCCTCCAGAGAGATGGGAAATCTAAGTCAAGATAGAGTACCCCAAGATGATGTT 693
 DB |||||||
 QY 3130 GGTCTCTTAGGATGATTTTGTACTATATGATTTAGGGGAAAAACACCAATTCAGCAGATAA 3189
 DB |||||||
 QY 692 GGTCTCTTAGGATGATTTTGTACTATATGATTTAGGGGAAAAACACCAATTCAGCAGATAA 633
 DB |||||||
 QY 3190 TTAATCAGATTTCTAAATTTACATGCAATATGATCTTAATCATGAATTTGAATTTCCG 3249
 DB |||||||
 QY 632 TTAATCAGATTTCTAAATTTACATGCAATATGATCTTAATCATGAATTTGAATTTCCG 573
 DB |||||||
 QY 3250 ATATTCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAGGGGACCCCAAAAC 3309
 DB |||||||
 QY 572 ATATTCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAGGGGACCCCAAAAC 513
 DB |||||||
 QY 3310 AGAGATATCCATTTCTGAGCTCTGAGCTCATCTATGTTTCAAAATTCAAATTCATCCAG 3369
 DB |||||||
 QY 512 AGAGATATCCATTTCTGAGCTCTGAGCTCATCTATGTTTCAAAATTCAAATTCATCCAG 453
 DB |||||||
 QY 3370 TTAACCAATGGCCAGGGAACCACTGAGAAATGAATATGTTCTGGGCCAATTTGTTG 3429
 DB |||||||
 QY 452 TTAACCAATGGCCAGGGAACCACTGAGAAATGAATATGTTCTGGGCCAATTTGTTG 393
 DB |||||||
 QY 3430 GTCTGAATTTCTCTAACTCCATTTTGAAGCTGCTAAAACTTTATATGAACATATAGTG 3489
 DB |||||||
 QY 392 GTCTGAATTTCTCTAACTCCATTTTGAAGCTGCTAAAACTTTATATGAACATATAGTG 333
 DB |||||||
 QY 3490 GTGGTGAAGTCAATATTTCTTCATCTCTCAAGACTTTTGAAGGAAAAAGGGAAAAAAT 3549
 DB |||||||
 QY 332 GTGGTGAAGTCAATATTTCTTCATCTCTCAAGACTTTTGAAGGAAAAAGGGAAAAAAT 273
 DB |||||||
 QY 3550 GATTTGACGTTATTCGTAATGTCAGATAGGAGTATAAAATATA---TTGACCTGTATA 3606
 DB |||||||
 QY 272 GATTTGACGTTATTCGTAATGTCAGATAGGAGTATAAAATATACTCTCGGCCCCCCCM 213
 DB |||||||
 QY 3607 CTCTTGAATCCCTGTGGAAATCTACATTTGAAGACCAATCACTCTGAAGTGTATTACG 3666
 DB |||||||
 QY 212 CYCTTSAATCCCTGTGGAAATCTACATTTGAARACAAATCACTCTGAAGTGTATTACG 153
 DB |||||||
 QY 3667 AAAAAAATTCAGTCAGATTTATCTTTTAAAGAAAACTGTAAATAAGCAACCACTATTGG 3726
 DB |||||||
 QY 152 AAAAAAATTCAGT-AGATTATCTTTTAAAGAAAACTGTAAATAAGCAACCACTATTGG 94
 DB |||||||
 QY 3727 CACTGTATATTTAGTACTGTTTCTGTTTATGCTCTTGTGTAATCTACTTGTACA 3786
 DB |||||||

```

Db      93 KACTGTATWTTTKTWAYTKTTTTTTTTTTTATAGTTTCTGTAACTTACTTGACW 34
QY      3787 TCATTTCCTCTTGGAAATAGTGGTGGATAGCA 3819
Db      33 TTTTTCCTTCTTGGAAATAGTGGTGGATAGCA 1

RESULT 9
AUI13100 LOCUS
DEFINITION AUI13100 NT2RP4 Homo sapiens cDNA clone NT2RP4001283 5', mRNA EST 01-AUG-2002
sequence.
ACCESSION AUI13100
VERSION AUI13100.1 GI:10993639
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 920)
AUTHORS Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J.,
Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuho,Y. and
Isogai,T.
HRI human cDNA project (Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y.,
Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T.,
Sugano,S., Masuho,Y., Isogai,T.)
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
LOCATION/Qualifiers
1..920
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP4001283"
/cell_type="teratocarcinoma"
/clone_lib="NT2"
/clone="NT2RP4"
/notes="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"

FEATURES
source
1..920

ORIGIN
Query Match 20.1%; Score 777.4; DB 9; Length 920;
Best Local Similarity 97.8%; Pred. No. 4.8e-154;
Matches 840; Conservative 0; Mismatches 13; Indels 6; Gaps 5;

QY 1150 TGGCAACAACCCAGAGACTGGTTGAGTTTGTCTCAAACTAGAGAAAAACAGTGTC 1209
Db 1 TGGCAACAACCCAGAGACTGGTTGAGTTTGTCTCAAACTAGAGAAAAACAGTGTC 60

QY 1210 CGCTAAGTGTGTCCTTTTAAATAAATTTGATGGTGTTCACGTCAAGCAATTTGAAGCGC 1269
Db 61 CGCTAAGTGTGTCCTTTTAAATAAATTTGATGGTGTTCACGTCAAGCAATTTGAAGCGC 120

QY 1270 TTCCCCCAGATAAATATGGCCAAATCAGAGTTTTCGTAGAAATTCAGTGAGATTTCGTC 1329
Db 121 TTCCCCCAGATAAATATGGCCAAATCAGAGTTTTCGTAGAAATTCAGTGAGATTTCGTC 180

QY 1330 AATTAAAGCTATTTCAGAGCCAGATGATGACGTGACTACTTCAAAATGCCAGAGCAA 1389
Db 181 AATTAAAGCTATTTCAGAGCCAGATGATGACGTGACTACTTCAAAATGCCAGAGCAA 240

QY 1390 ACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAATTTGACTGTCACAGGTA 1449

```

```

Db      241 ACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCAAAATTTGAATGTCACAGGTA 300
QY      1450 ATGTCAAAAAAGTAACAACCTTCTTCAAAAAGCTGTAGACGTGGAGCAGTACACTAG 1509
Db      301 ATGTCAAAAAAGTAACAACCTTCTTCAAAAAGCTGTAGACGTGGAGCAGTACACTAG 360
QY      1510 AAATGCTGGAATTCCTCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGCTTTTCAGAGG 1569
Db      361 AAATGCTGGAATTCCTCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGCTTTTCAGAGG 420
QY      1570 AGAAAAAGAAATTTATCAGCATCTACGATTAATTAACCTGCCCAAGAAATCATTTCCGGTT 1629
Db      421 AGAAAAAGAAATTTATCAGCATCTACGATTAATTAACCTGCCCAAGAAATCATTTCCGGTT 480
QY      1630 CACTTGGGCATTTACAGAAATAGGAACAACAGCTTGTGATTCAGAGGACAGACTACTAAAG 1689
Db      481 CACTTGGGCATTTACAGAAATAGGAACAACAGCTTGTGATTCAGAGGACAGACTACTAAAG 540
QY      1690 CCAGTTTTTTATTCGAGAGAACATGCCACCAAGATGCAGAAATAGTTTACCGGAATT 1749
Db      541 CCAGTTTTTTATTCGAGAGAACATGCCACCAAGATGCAGAAATAGTTTACCGGAATT 600
QY      1750 CATTGAGCAAACTAAACAACACTAAACAGCATGCCCATTTGGAGAGTCCCGATTAAAC 1809
Db      601 CATTGAGCAAACTAAACAACACTAAACAGCATGCCCATTTGGAGAGTCCCGATTAAAC 660
QY      1810 TTCTAAATAGCCAGATTTGTGATGTGAAGACAGATGATTCAGTTGTACCTTTGTTTATGA 1869
Db      661 TTCTAAATAGCCAGATTTGTGATGTGAAGACAGATGATTCAGTTGTACCTTTGTTTATGA 719
QY      1870 AAAGCAAACTCTAGATCAGAAATG-CGAGATTTGTTGCTGCTGATCTAAACCAAGT 1928
Db      720 AAAGCAAACTCTAGATCAGAAATG-CGAGATTTGTTGCTGCTGATCTAAACCAAGT 779
QY      1929 GGAATGATTTCTGTGAATTAAGAAATTTAAAGCTGTCTCAAAATAGTCATTTTCAAGGAA 1988
Db      780 GGAATGATTTCTGTGAATTAAGAAATTTAAAGCTGTCTCAAAATAGTCATTTTCAAGGAA 836
QY      1989 CCTCTGGTGTGATGAGAA 2007
Db      837 CCT-TGGTGTGATGAGAA 854

RESULT 10
AUI19487 LOCUS
DEFINITION AUI19487 HEMBAI Homo sapiens cDNA clone HEMBAI005932 5', mRNA
sequence.
ACCESSION AUI19487
VERSION AUI19487.1 GI:10934722
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 837)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
HRI human cDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

```

```
FEATURES
  source      Location/Qualifiers
1..837
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="HEMBA1005932"
  /tissue_type="whole embryo, mainly head"
  /dev_stage="embryo, 10 weeks"
  /clone_lib="HEMBA1"
  /note="Vector: pME18SFL3"

ORIGIN
Query Match      19.9%; Score 771; DB 9; Length 837;
Best Local Similarity 96.0%; Pred. No. 1.1e-152;
Matches 800; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 897 GTGAATGTAACTGTACAGACTTCCTAGAAAACAGTTGGTGGTCCATCTTTCAATTC 956
Db 5 GCGACGGGAAATCAACAGCTGTTCGGGAAAGAGTTGGTGGTCCATCTTTCAATTC 64
QY 957 CCAGTGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAAATGACAAT 1016
Db 65 CCAGCGAGCTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAAATGACAAT 124
QY 1017 GATTCCATATGAACAAAGTGAAGACATTTAAATAAGTTTAAATAAGTGAAGACCTTACT 1076
Db 125 GATTCCATATGAACAAAGTGAAGACATTTAAATAAGTTTAAATAAGTGAAGACCTTACT 184
QY 1077 GATGAATCAAGCTTGAATAAATTTCTGTGATACAGATACTCGGAACCTGTTAAC 1136
Db 185 GATGAATCAAGCTTGAATAAATTTCTGTGATACAGATACTCGGAACCTGTTAAC 244
QY 1137 CAATATGATGATGGCAAAACACCCAGAGGACTGGTGTAGTTTGGTCTCAAACTAGAG 1196
Db 245 CAATATGATGATGGCAAAACACCCAGAGGACTGGTGTAGTTTGGTCTCAAACTAGAG 304
QY 1197 ABAACAGTGTCCGCTAAGTGATGCTCTTTTAAATAATGATGTGCTGTACAGTCAA 1256
Db 305 ABAACAGTGTCCGCTAAGTGATGCTCTTTTAAATAATGATGTGCTGTACAGTCAA 364
QY 1257 GCAATTTGAAGCGCTTCCCCAGATAAATATGCCCCAAATGAGATTTTGTGTAGAAATCAA 1316
Db 365 GCAATTTGAAGCGCTTCCCCAGATAAATATGCCCCAAATGAGATTTTGTGTAGAAATCAA 424
QY 1317 GTGAGATTTGTGAATTAAGCTATTCAAGAGCCAGATGATGACGTGACTACTTTCAA 1376
Db 425 GTGAGATTTGTGAATTAAGCTATTCAAGAGCCAGATGATGACGTGACTACTTTCAA 484
QY 1377 ATGCCAGAGCAAACTGCAAGAAATTTGCTTTTGTATATATCTTTTGCACAAATTTGAA 1436
Db 485 ATGCCAGAGCAAACTGCAAGAAATTTGCTTTTGTATATATCTTTTGCACAAATTTGAA 544
QY 1437 CTGTCAAGTAAATGTCAAAAAAGTAAACAACTTCTTCAAAAGCTGTAGAACGTGGA 1496
Db 545 CTGTCAAGTAAATGTCAAAAAAGTAAACAACTTCTTCAAAAGCTGTAGAACGTGGA 604
QY 1497 GCAGTACCACATAGAAATCTGGAATTCCTCGGGAATTTAACTTCCAAAAGAGCAG 1556
Db 605 GCAGTACCACATAGAAATCTGGAATTCCTCGGGAATTTAACTTCCAAAAGAGCAG 664
QY 1557 CTGCTTTTCAGAGGAGGAAAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAA 1616
Db 665 CTGCTTTTCAGAGGAGGAAAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAA 724
QY 1617 TCATTTTCGGTTCACTTGGSCATTTACAGATAGGACACACAGTTGTGATTCAGAGGA 1676
Db 725 TCATTTTCGGT-TCACTTGGSCATTTACAGATAGGACACACCGTTGTGATTTACAGAGGA 783
QY 1677 CAGACTACTAAAGCCAGGTTTTATATGAGAGAGAACATGCCACCACAAGATGC 1729
Db 784 CAGACTACTTAAGCCAGGTTTTATATGAGAGAGAACATGCCCCCCCANGATGC 836

LOCUS      BI462468
DEFINITION 603203960F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5269695 5', mRNA sequence.
ACCESSION  BI462468
VERSION    BI462468.1 GI:15253124
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 878)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11680 row: h column: 16
            High quality sequence stop: 819.

FEATURES
  source      Location/Qualifiers
1..878
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:5269695"
  /lab_host="DH10B"
  /clone_lib="NIH MGC 97"
  /note="Organ: testis; Vector: pBluescriptR (modified
  pluescript KS+); Site: 1: BamHI; Site: 2: SalI-XhoI
  (gtcag); Oligo-dT primed using primer
  5'-TTTTTTTTTTTTTTVN-3', size-selected for
  insert size 2.2 kb and normalized to R0T 5. This is a
  primary library enriched for full-length clones and
  constructed using the Cap-trapper method (Carninci, in
  preparation). Library constructed by M. Brownstein
  (NIH/NHGRI, National Institutes of Health). Note: this is
  a NIH_MGC Library."

ORIGIN
Query Match      19.7%; Score 763.2; DB 12; Length 878;
Best Local Similarity 95.5%; Pred. No. 4.9e-151;
Matches 829; Conservative 0; Mismatches 33; Indels 6; Gaps 4;

QY 951 ATTTCCCGAGTCAGTTTCTCTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAAATG 1010
Db 6 ATTTCCCGAGTCAGTTTCTCTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAAATG 65
QY 1011 ACAATGTATCCATATGACAAAGTGAAGAGACATTAATAAGTTTAAATGAAGAC 1070
Db 66 ACAATGTATCCATATGACAAAGTGAAGAGACATTAATAAGTTTAAATGAAGAC 125
QY 1071 CTTACTGTGAATGAAGCTTGAATTAATAATTTCTGTGATCTACAGATAACTCGGAACT 1130
Db 126 CTTACTGTGAATGAAGCTTGAATTAATAATTTCTGTGATCTACAGATAACTCGGAACT 185
QY 1131 GTTAACCAATTTATGATGATGCAAAACCCAGAGGACTGGTGGTGTGCTCAA 1190
Db 186 GTTAACCAATTTATGATGATGCAAAACCCAGAGGACTGGTGGTGTGCTCAA 245
QY 1191 CTAGAGAAAACAGTGTCCGCTAAGTGATGCTCTTTTAAATAAATGATGTGCTTAC 1250
Db 246 CTAGAGAAAACAGTGTCCGCTAAGTGATGCTCTTTTAAATAAATGATGTGCTTAC 305
QY 1251 AGTCAAGCAATTTGAAGCGCTTCCCGAGATAAATAATGCGCAAAATGAGATTTTCTAGA 1310
```

```
Db 306 AGTCAAGCAATTGAAGCGCTTCCCCAGATAAATATGCCCCAAATGAGAGTTTGTGTAGA 365
Qy 1311 ATTCAAGTGAATTTGCTGAATTAAGAACTATTCAAGAGCCAGATGATGACGTGACTAC 1370
Db 366 ATTCAAGTGAATTTGCTGAATTAAGAACTATTCAAGAGCCAGATGATGACGTGACTAC 425
Qy 1371 TTTCAAATGGCCAGAGCAAACTGCAAGAAATTTGCTTTTCTTTCATATATCTTTTGCACAA 1430
Db 426 TTTCAAATGGCCAGAGCAAACTGCAAGAAATTTGCTTTTCTTTCATATATCTTTTGCACAA 485
Qy 1431 TTTCAAATGTCACAAAGTAAATGTCACAAAGTAAATGTCACAAAGTAAATGTCACAAAGT 1490
Db 486 TTTCAAATGTCACAAAGTAAATGTCACAAAGTAAATGTCACAAAGTAAATGTCACAAAGT 545
Qy 1491 CGTGGAGCAGTACCTAGAAATGCTGGAATTCCTGCGGAAATTTAAACCTCCAAAAA 1550
Db 546 CGTGGAGCAGTACCTAGAAATGCTGGAATTCCTGCGGAAATTTAAACCTCCAAAAA 605
Qy 1551 AAGCAGCTGCTTTC--AGAGGAGGAAAGAGAAATTTATCAGCATCTACGTTATTAAC 1607
Db 606 AAGCAGCTGCTTTCAGAGAGGAGGAAAGAGAAATTTATCAGCATCTACGTTATTAAC 665
Qy 1608 GCCCAAGAACTATTTCCGGTTCACTTGGGCATTTACAGAAATAGGAACAAACAGTTGTGAT 1667
Db 666 GCCCAAGAACTATTTCCGGTTCACTTGGGCATTTACAGAAATAGGAACAAACAGTTGTGAT 724
Qy 1668 TCCAGAGGACAGACTATTAAGCCAGGTTTTTATATGAGAGACATGCCACCAAGAT 1727
Db 725 TCCAGAGGACAGACTATTAAGCCAGGTTTTTATATGAGAGACATGCCACCAAGAT 783
Qy 1728 GCAGAAATAGTTTACCGGAATTCATTGAGACAAACTTAAACAGTATGCCCA 1787
Db 784 GC-CAATAGTTTACCGGAATTCATTGAGACAAACTTAAACAGTATGCCCA 842
Qy 1788 TTGGAAGAGTCCAGTTTACCTCTAA 1815
Db 843 TTGGAAGAGTCCAGTTTACCTCTAA 870

RESULT 12
AA744546 912 bp mRNA linear EST 07-FEB-1998
LOCUS ny79c06.s1 NCI_CGAP GCBI Homo sapiens cDNA clone IMAGE:1284490 3'
DEFINITION similar to gb:M86699 DUAL SPECIFICITY PROTEIN KINASE TTK (HUMAN);
mRNA sequence.
ACCESSION AA744546
VERSION AA744546.1 GI:2783310
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (Bases 1 to 912)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1024 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 479.
Location/Qualifiers
```

```
source
1. 912
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1284490"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/clone_lib="NCI_CGAP GCBI"
/note="vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-,
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGGGCGGCTCATTTTCTTTTCTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
```

Query Match	19.1%	Score 737.2	DB 9	Length 912
Best Local Similarity	93.3%	Pred. No. 1.6e-145		
Matches 836	Conservative 0	Mismatches 53	Indels 7	Gaps 6
Qy 2971	TTGGGATTGCAACCAATGCAACAGATACCAACAGTGTGTTTAAAGATTCTCAGGTG	3030		
Db 899	TTTGGGTCTCCACCCACTGCCCCCGGATCCAC-AGTGTGTTAAGATTCCTCGGTG	841		
Qy 3031	GCAGATTAATATATATGCAACAGAGCAATCAAGATATGTCTTCTCCAGAGAGATG	3090		
Db 840	CCACGTTTAAAT-TTAGCCCAACCGAGACAACCAAGGATATGTCTTCT-CACTGAGAATG	784		
Qy 3091	GGAAATCTAAGTCAAGATAGCCCCCAAGTGTGTTGGTCTTAGGATGATTTTGT	3150		
Db 783	GGAACTCTAAGTC-AAGATAAGCCCCCAAGTGTGTTGGTCTTAGGATGATTTTGT	725		
Qy 3151	ACTATATGACTTACGGGAAAAACACATTTCCAGCAGATAAATTAATCAGATTCTTAAATAC	3210		
Db 724	ACTATATGACTTACGGGAAAAACACATTTCCAGCAGATAAATTAATCAGATTCTTAAATAC	665		
Qy 3211	ATGCCATAATGATCCTTAATCATGAATTAATTCGGATATTCAGAGAGAGATCTTC	3270		
Db 664	ATGCCATAATGATCCTTAATCATGAATTAATTCGGATATTCAGAGAGAGATCTTC	605		
Qy 3271	AAGATGCTGTAAAGTGTGTTTAAAGGGGACCCCAACAGAGGATATCCATTCCTGAGC	3330		
Db 604	AAGATGCTGTAAAGTG-TGTTTAAAGGGGACCCCAACAGAGGATATCCATTCCTGAGC	546		
Qy 3331	TCCTGGCTCATCCATATGTTCAAAATCAAACTCATCCAGTTAACCAATGGCCAAAGGAA	3390		
Db 545	TCCTGGCTCATCCATATGTTCAAAATCAAACTCATCCAGTTAACCAATGGCCAAAGGAA	486		
Qy 3391	CCACTGAGAAATGAATATGTTCTGGCCCAACTGTTGGTCTCAATTCCTCAACTCCA	3450		
Db 485	CCACTGAGAAATGAATATGTTCTGGCCCAACTGTTGGTCTCAATTCCTCAACTCCA	426		
Qy 3451	TTTTGAAAGCTGCTAAAACTTTATATGAACACTAGTGGTGGTGAAGTCAATATCTT	3510		
Db 425	TTTTGAAAGCTGCTAAAACTTTATATGAACACTAGTGGTGGTGAAGTCAATATCTT	366		
Qy 3511	CATCTCCCAAGACTTTGAAAAAAGGGGAAAAAATGATTTCGAGTTATTCGTAAATG	3570		
Db 365	CATCTCCCAAGACTTTGAAAAAAGGGGAAAAAATGATTTCGAGTTATTCGTAAATG	306		
Qy 3571	TCAGATAGAGGTATAAATAATATTGGACTGTTTACTCTTGAATCCCTGTGAAATCTA	3630		
Db 305	TCAGATAGAGGTATAAATAATATTGGACTGTTTACTCTTGAATCCCTGTGAAATCTA	246		
Qy 3631	CATTGAGACAACTACTCTGAAGTGTATCAGCAAAAAAATTCAGTGAGATTAATCT	3690		
Db				

```

Db      245  CATTGAGACAACATCACTCTGAAGTGTATATCAGCAAAAAAATTCAGT-AGATTATCT 187
Qy      3691  TTAAGAGAAACTGTAAATAATAGCAACCACTATGGCACTGTATATATTTAGACTTGTT 3750
Db      186  TTAAGAGAAACTGTAAATAATAGCAACCACTATGGCACTGTATATTTAGACTTGTT 127
Qy      3751  TTCCTGTTTTATCTCTGTTGTAATCTACTACTGATCATCATTTTACTCTTGAATAGTGG 3810
Db      126  TTCCTGTTTTATCTCTGTTGTAATCTACTACTGATCATCATTTTACTCTTGAATAGTGG 67
Qy      3811  TGGATAGCACTATATCTAAACAACTTGTAAATAAGTTTGTGCTAAATGA 3866
Db      66  TGGATAGCACTATATCTAAACAACTTGTAAATAAGTTTGTGCTAAATGA 11

RESULT 13
BO428904
LOCUS      792 bp mRNA linear EST 24-MAY-2002
DEFINITION AGENCOURT_7905694 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6105546
5', mRNA sequence.
ACCESSION BO428904
VERSION BO428904.1 GI:21167980
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 792)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2345 row: c column: 19
High quality sequence stop: 548.
FEATURES
Location/Qualifiers
1..792
/organism="Homo sapiens"
/mo_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6105546"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 82"
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggccattatggcc); Site 2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGCGCATTTAGCC-3' and 3' adaptor sequence:
5'-ATTGTAGAGCGCGGCGGCACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
```

ORIGIN

```

Query Match      18.9%; Score 731.4; DB 13; Length 792;
Best Local Similarity 98.6%; Pred. No. 2.7e-144;
Matches 768; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

Qy      1995  GTGTCAGATGAAGAGTTCGACTTATTTACTGATTCATTAACCCCTGAAGATAAA 2054
Db      1  GGGTCAGATGAAGAGTTCGAACTTATTTACTGATTCATTAACCCCTGAAGATAAA 60
Qy      2055  ACGGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAGTATCAAGAACCAAGGTT 2114
Db      61  ACGGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAGTATCAAGAACCAAGGTT 120
```

```

Qy      2115  CCAGAGAGTAACACAGAAACAGTGGCAAGCTAAGAGAAAGTCCAGAGTGATTAAACAGAA 2174
Db      121  CCAGAGAGTAACACAGAAACAGTGGCAATCTAAGAGAAAGTCCAGAGTGATTAAACAGAA 180
Qy      2175  CTTGCTGCACTCTTCAAAATCAGTGGCAGATTCGGAGTGTAGCCGAAAAGTTAATACAGAG 2234
Db      181  CTTGCTGCACTCTTCAAAATCAGTGGCAGATTCGGAGTGTAGCCGAAAAGTTAATACAGAG 240
Qy      2235  CAGAAACATACACACTTTTGAGCAACCTGTCTTTTTCAGTTTCAAAACAGTCCACCAATA 2294
Db      241  CAGAAACATACACACTTTTGAGCAACCTGTCTTTTTCAGTTTCAAAACAGTCCACCAATA 300
Qy      2295  TCAACATCTAAATGGTTTGACCCCAAAATCTATTTTGAAGACACCAGCAGCAATACCTTG 2354
Db      301  TCAACATCTAAATGGTTTGACCCCAAAATCTATTTTGAAGACACCAGCAGCAATACCTTG 360
Qy      2355  GATGATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAAGAAATGACTTCCACCTGCTGT 2414
Db      361  GATGATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAAGAAATGACTTCCACCTGCTGT 420
Qy      2415  CAGTTGCTCAACACCTTTATGGCCAAACCTGCTCTTTTCCAGCAGCAACAGCATCAAAATCT 2474
Db      421  CAGTTGCTCAACACCTTTATGGCCAAACCTGCTCTTTTCCAGCAGCAACAGCATCAAAATCT 480
Qy      2475  GCACCTCCACTTCAAAATTTACAGGTTTTCAGCATCTCTTTCAGCAAAATGAATGCATTCG 2534
Db      481  GCACCTCCACTTCAAAATTTACAGGTTTTCAGCATCTCTTTCAGCAAAATGAATGCATTCG 540
Qy      2535  GTTAAAGGAAGAATTTATTCCTATATTAAAGCAGATAGGAAGTGGAGGTTCAAGCAAGGTA 2594
Db      541  GTTAAAGGAAGAATTTATTCCTATATTAAAGCAGATAGGAAGTGGAGGTTCAAGCAAGGTA 600
Qy      2595  TTTGAGGTGTTAAATGAAAGAAACAGATATATGCTATATAATATGTGAACCTTAGAAGAA 2654
Db      601  TTTGAGGTGTTAAATGAAAGAAACAGATATATGCTATATAATATGTGAACCTTAGAAGAA 660
Qy      2655  GCAGATAACCAACTCTTCATAGTTA-CGGAGCAAGAAATAGCTTATTTCGAATAAACTACA 2713
Db      661  GCAGATAACCAACTCTTCATAGTTA-CGGAGCAAGAAATAGCTTATTTCGAATAAACTACA 720
Qy      2714  ACAACACAGT-GATAAGATCAT-CCGACTTTATGATTATGAATCAACGACCAAGTACAT 2770
Db      721  ACAACNCACTGGATAAGATCATCCCGACTTTATGATTATGAATCAACGACCAAGTACCT 779

RESULT 14
BG718184
LOCUS      602696148F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4828282 5',
DEFINITION mRNA sequence.
ACCESSION BG718184
VERSION BG718184.1 GI:13997371
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 778)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL10745 row: p column: 11
```

High quality sequence stop: 770.
 Location/Qualifiers
 1..778
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4828282"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC 97"
 /note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.2 Kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 18.7%; Score 723.4; DB 12; Length 778;
 Best Local Similarity 98.8%; Pred. No. 1.3e-142;
 Matches 760; Conservative 0; Mismatches 6; Indels 3; Gaps 3;
 931 AGTTGGGTTCCATCTTTTCATTTCCCGAGTGCAGTTTCTGTAGAAATCGAATCCGAGG 990
 Db |||||
 5 AGTTGGGTTCCATCTTTTCATTTCCCGAGCGCAGCTTCTGTAGAAATCGAATCCGAGG 64
 |||||
 991 ATTTAAGTGGCAGAGAATTGCAATTCATTAATGAACAAAGTGAGAGACATTAATAA 1050
 Db |||||
 65 ATTTAAGTGGCAGAGAATTGCAATTCATTAATGAACAAAGTGAGAGACATTAATAA 124
 |||||
 1051 ATAAAGTTAAATGAGACCTTACTGATGACTTAAGCTTGAATAAAATTCGCTGATA 1110
 Db |||||
 125 ATAAAGTTAAATGAGACCTTACTGATGACTTAAGCTTGAATAAAATTCGCTGATA 184
 |||||
 1111 CTACAGATACTCGGGAACCTGTTAAACCAATTTATGATGATGGCAACCAACCCAGAGACT 1170
 Db |||||
 185 CTACAGATACTCGGGAACCTGTTAAACCAATTTATGATGATGGCAACCAACCCAGAGACT 244
 |||||
 1171 GGTGAGTTTGTCTCAAACTAGAGAAAACAGTGTTCGCTTAAGTATGCTCTTTTAA 1230
 Db |||||
 245 GGTGAGTTTGTCTCAAACTAGAGAAAACAGTGTTCGCTTAAGTATGCTCTTTTAA 304
 |||||
 1231 ATAAATGATTTGTCGTTACAGTCAAGCAATTTGAAGCGCTTCCCCAGATAAATATGGCC 1290
 Db |||||
 305 ATAAATGATTTGTCGTTACAGTCAAGCAATTTGAAGCGCTTCCCCAGATAAATATGGCC 364
 |||||
 1291 AAAATGAGAGTTTGTCTAGAAATTCAGTGAATTTGCTGAATTAAGAGCTATTCAAGAGC 1350
 Db |||||
 365 AAAATGAGAGTTTGTCTAGAAATTCAGTGAATTTGCTGAATTAAGAGCTATTCAAGAGC 424
 |||||
 1351 CAGATGATGACGTCGCTACTTTCAATGCGCAGACAACTGCAAGAAATTTGCTTTTG 1410
 Db |||||
 425 CAGATGATGACGTCGCTACTTTCAATGCGCAGACAACTGCAAGAAATTTGCTTTTG 484
 |||||
 1411 TTCTATATCTTTTGCACAAATTTGAACTGTCAAGGTAATGTCAAA-AAAAAGTAAACAA 1469
 Db |||||
 485 TTCTATATCTTTTGCACAAATTTGAACTGTCAAGGTAATGTCAAAAGTAAACAA 544
 |||||
 1470 CTTCTTCAAAAGCTGTAGAACGTGGAGAGTACCCTAGAAATGCTGGAATTTGCCCTG 1529
 Db |||||
 545 CTTCTTCAAAAGCTGTAGAACGTGGAGAGTACCCTAGAAATGCTGGAATTTG-CCCTG 603
 |||||
 1530 CGGAATTTAAACCTCCAAAAGAACGCTGTTTTCAGAGGAGAA-AAAGAAATTTATC 1588
 Db |||||
 604 CGGAATTTAAACCTCCAAAAGAACGCTGTTTTCAGAGGAGAACAGAGAGATCTATC 663
 |||||
 1589 AGCATCTACGGTATTACTGCCCAAGAAATCATTTTCGGTTCATCTGGGCAATTTACAGAA 1648
 Db |||||
 664 AGCATCTACGGTATTACTGCCCAAGAAATCATTTTCGGTTCATCTGGGCAATTTACAGAA 723
 |||||

QY 1649 TAGGAACAAACAGTTGCTGATTCAGAGGACAGACTACTAAACCCAGGTTT 1697
 |||||
 Db 724 TAGGAACAAACAGTTGCTGATTCAGAGGACAGACTACTAAACCCAGGTTT 772
 |||||

RESULT 15
 AUI32533
 LOCUS AUI32533 NT2RP4 Homo sapiens cDNA clone NT2RP4000021 5', mRNA
 DEFINITION sequence.
 ACCESSION AUI32533
 VERSION AUI32533.1 GI:10992887
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 821)
 AUTHORS Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J.,
 Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuhio,Y. and
 Isogai,T.
 TITLE HRI human cDNA project (Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y.,
 Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T.,
 Sugano,S., Masuhio,Y., Isogai,T.)
 JOURNAL Unpublished (2000)
 COMMENT Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project, 5'- & 3'-end one pass sequencing: Helix
 Research Institute; cDNA library construction: Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.
 FEATURES
 Location/Qualifiers
 1..821
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NT2RP4000021"
 /cell_type="teratocarcinoma"
 /cell_line="NT2"
 /clone_lib="NT2RP4"
 /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
 cells after 2-weeks retinoic acid (RA) induction"

ORIGIN
 Query Match 18.6%; Score 720.6; DB 9; Length 821;
 Best Local Similarity 95.1%; Pred. No. 5.2e-142;
 Matches 774; Conservative 0; Mismatches 37; Indels 3; Gaps 3;
 897 GTGAATGTAAAGCTGTACAGACTTCCTAGAAAACAGTTGGTTCCTATTTTCATTTCC 956
 Db 5 GCGAGCGGGAATTAACAAACGTTTTCGGAAGAGAGTTGGTTCCTATTTTCATTTCC 64
 |||||
 957 CCAGTGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGCAGAGAAATTCACAATT 1016
 Db 65 CCAGCGCAGCTTCTGTAGAAATGGAATCCGAGGATTTAAGTGCAGAGAAATTCACAATT 124
 |||||
 1017 GATTCCTAATGAACAAAGTGAGAGACATTAATAATAGTTTAAAAATGAAGACCTTACT 1076
 Db 125 GATTCCTAATGAACAAAGTGAGAGACATTAATAATAGTTTAAAAATGAAGACCTTACT 184
 |||||
 1077 GATGAACCTAAGCTTCAATATAAATTTCTGCTGATACTACAGATACTCGGAACCTGTTAAC 1136
 Db 185 GATGAACCTAAGCTTCAATATAAATTTCTGCTGATACTACAGATACTCGGAACCTGTTAAC 244
 |||||
 1137 CAATATGATGATGGCAACCAACCCAGAGAGCTGGTTGAGTTTGTGCTCAAACTAGAG 1196
 Db 245 CAATATGATGATGGCAACCAACCCAGAGAGCTGGTTGAGTTTGTGCTCAAACTAGAG 304
 |||||

Qy 1197 AAAACAGTGTTCGCTAGTGATGCTCTTTTAAATAAATTTGTTGGTTACAGTCAA 1256
Db |||||
Qy 305 AAAACAGTGTTCGCTAGTGATGCTCTTTTAAATAAATTTGTTGGTTACAGTCAA 364
Db |||||
Qy 1257 GCAATTGAAGCGCTTCCCCAGATPAATAATGGCCAAAATGAGAGTTTGTAGAAATCAA 1316
Db |||||
Qy 365 GCAATTGAAGCGCTTCCCCAGATPAATAATGGCCAAAATGAGAGTTTGTAGAAATCAA 424
Db |||||
Qy 1317 GTGAGATTGCTGAATTAAGCTATTCAAGAGCCGATGATGACGTGACTACTTTCAA 1376
Db |||||
Qy 425 GTGAGATTGCTGAATTAAGCTATTCAAGAGCCGATGATGACGTGACTACTTTCAA 484
Db |||||
Qy 1377 ATGCCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGACAAATTTGAA 1436
Db |||||
Qy 485 ATGCCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGACAAATTTGAA 544
Db |||||
Qy 1437 CTGTACAGAGTAATGTCMAAAAGTAACAACTTCTTCAAAAAGCTGTAGAACGTGA 1496
Db |||||
Qy 545 CTGTACAGAGTAATGTCMAAAAGTAACAACTTCTTCAAAAAGCTGTAGAACGTGA 604
Db |||||
Qy 1497 GCAGTACCCTAGAAATGCTGGAATTTGCCCTGCGGAATTTAAACCTCCAAAAGACAG 1556
Db |||||
Qy 605 GCAGTACCCTAGAAATGCTGGAATTTGCCCTGCGGAATTTAAACCTCCAAAAGACAG 664
Db |||||
Qy 1557 CTGCTTTACAGAGGAGAAAGAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAA 1616
Db |||||
Qy 665 CTGCTTTACAGAGGAGAAAGAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAA 724
Db |||||
Qy 1617 TCATTTCCGGTTCACCTGGGCATTTACAGAATAGGAACAACAGTTTGTGATTCAGAGGA 1676
Db |||||
Qy 725 TCATTTCCGG-TCACTTGGCATTT-CAGATAGG-ACCACAGTTGTGATTCAGANGA 781
Db |||||
Qy 1677 CAGACTACTAAAGCCAGGTTTTTATATGGAGAGA 1710
Db |||||
Qy 782 CAGACTCTAAAGNCAGGTTTTATATGGAGAGACA 815
Db |||||

Search completed: September 30, 2004, 08:58:10
Job time : 8594 secs

This Page Blank (uspto)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	3866	100.0	3866	6	AB1677956	Ovary can
2	3866	100.0	3866	6	ABV73974	Human tyr
3	3866	100.0	3866	8	ACA62263	Human cDN
4	3866	100.0	3866	9	ADD49331	Human lun
5	3866	100.0	3866	9	ADE40436	Human TTK
6	3866	100.0	3866	9	AD838368	Human pro
7	3866	100.0	3866	9	AD831644	Human 168
8	2909.12	75.3	2975	9	ADC37170	Nuclear f
9	2906	75.2	2980	4	AAS03025	Human dia
10	2569.2	66.5	2574	7	AB76453	Lung can
11	2569.12	66.5	2574	9	ADB80469	Ovarian c
12	1633	42.2	2860	9	ADC37168	Nuclear f
13	395.56	10.2	32404	8	ADA02894	Human BLR
14	395.56	10.2	32404	9	AD872632	Human BLR
15	395.56	10.2	32404	9	ADC85373	Mouse Blr
16	390.4	10.1	57130	7	ABQ77243	Human MAR
17	386	10.0	122748	6	ABT10719	Human bre
18	385.6	10.0	174424	6	ABJ68122	Ovary can
19	385.4	10.0	949	4	AAR18360	Human cDN
20	385.4	10.0	32190	4	AAK89112	Human dig
21	385.4	10.0	32190	5	AAS31862	Human liv
22	385.4	10.0	32190	6	ABN90217	Human liv
23	385.2	10.0	207433	5	ABY72040	Gene 216

QY 2881 TTCCACAAATCCATCAACATGGCATTGTTTACAGTGATCTTAAACACAGCTAACTTTCTGA 2940
 Db 2881 TTCCACAAATCCATCAACATGGCATTGTTTACAGTGATCTTAAACACAGCTAACTTTCTGA 2940
 QY 2941 TAGTTGATGAATGCTAAAGCTAAATGATTTGGGATTGCAAAACCAAATGCAACACAGATA 3000
 Db 2941 TAGTTGATGAATGCTAAAGCTAAATGATTTGGGATTGCAAAACCAAATGCAACACAGATA 3000
 QY 3001 CAACAAGTGTGTTAAAGATTTCTCAGGTGGCAGATTAATATATATGCCCACCAAGACAA 3060
 Db 3001 CAACAAGTGTGTTAAAGATTTCTCAGGTGGCAGATTAATATATATGCCCACCAAGACAA 3060
 QY 3061 TCAAGATATGCTTTCTCCAGAGAGAAATCTTAAGTCAAAAGATAAGCCCCAAAA 3120
 Db 3061 TCAAGATATGCTTTCTCCAGAGAGAAATCTTAAGTCAAAAGATAAGCCCCAAAA 3120
 QY 3121 GTGATGTTTGGTCTTATAGGATGTAATTTGTACTATATATGATTCACGGGAAACACCAATTC 3180
 Db 3121 GTGATGTTTGGTCTTATAGGATGTAATTTGTACTATATATGATTCACGGGAAACACCAATTC 3180
 QY 3181 AGCAGATAATTAATCAGATTTCTAAATTTACATGCCATAATTTGATCCTTAATCATGAATTCG 3240
 Db 3181 AGCAGATAATTAATCAGATTTCTAAATTTACATGCCATAATTTGATCCTTAATCATGAATTCG 3240
 QY 3241 AATTTCCTCGATATCCAGAGAAAGATCTTCAAGATGTTGTTAAAGTGTGTTTAAAAAGGG 3300
 Db 3241 AATTTCCTCGATATCCAGAGAAAGATCTTCAAGATGTTGTTAAAGTGTGTTTAAAAAGGG 3300
 QY 3301 ACCCAAAACAGAGATATCCATTCCTGAGCTCTGGCTCATCATATGTTCAAAATTCAAA 3360
 Db 3301 ACCCAAAACAGAGATATCCATTCCTGAGCTCTGGCTCATCATATGTTCAAAATTCAAA 3360
 QY 3361 CTCATCCAGTTAAACCAATGGCCAAAGGAACCACTGAAGAAATGAATATGTTCTGGGCC 3420
 Db 3361 CTCATCCAGTTAAACCAATGGCCAAAGGAACCACTGAAGAAATGAATATGTTCTGGGCC 3420
 QY 3421 AACTTGTGTTGTCGAATTCCTTAACCTCCATTTTGAAGCTGTAACACTTTATATGAAC 3480
 Db 3421 AACTTGTGTTGTCGAATTCCTTAACCTCCATTTTGAAGCTGTAACACTTTATATGAAC 3480
 QY 3481 ACTATAGTGGTGGGAAAGTCATAATTTCTCATCTCCAGAGCTTTTGAAGAAAGGG 3540
 Db 3481 ACTATAGTGGTGGGAAAGTCATAATTTCTCATCTCCAGAGCTTTTGAAGAAAGGG 3540
 QY 3541 GAAAAAATGATTTGCGAGTTATTTCGTAATGTCAGATAGGAGGTATAAATATATTTGGACT 3600
 Db 3541 GAAAAAATGATTTGCGAGTTATTTCGTAATGTCAGATAGGAGGTATAAATATATTTGGACT 3600
 QY 3601 GTTATACCTTGATCCCTGCGAATCTACATTTGAAGACACATCCTCCTCAAGTGT 3660
 Db 3601 GTTATACCTTGATCCCTGCGAATCTACATTTGAAGACACATCCTCCTCAAGTGT 3660
 QY 3661 ATCAGCAAAAAAATTCAGTGAGATATCTTTAAAGAAACCTGTAAGATAGCAACAC 3720
 Db 3661 ATCAGCAAAAAAATTCAGTGAGATATCTTTAAAGAAACCTGTAAGATAGCAACAC 3720
 QY 3721 TTATGGCAGCTATATATTTAGACTGTTTCTCTGTTTATGCTCTGTTGTAATCTAC 3780
 Db 3721 TTATGGCAGCTATATATTTAGACTGTTTCTCTGTTTATGCTCTGTTGTAATCTAC 3780
 QY 3781 TTGACATCAATTTTACTCTTGGATAGTGGGTGGATAGCAAGTATATTTAAAAAATTTTG 3840
 Db 3781 TTGACATCAATTTTACTCTTGGATAGTGGGTGGATAGCAAGTATATTTAAAAAATTTTG 3840
 QY 3841 TAAATAAAGTTTGTGGCTAAATGA 3866
 Db 3841 TAAATAAAGTTTGTGGCTAAATGA 3866

RESULT 3

ID ACA62263 standard; cDNA; 3866 BP.
 XX

AC ACA62263;
 XX 11-AUG-2003 (first entry)
 DT Human cDNA encoding Tyrosine threonine kinase, TTK.
 DE Human; ss: gene; tyrosine threonine kinase; TTK; cancer; cytostatic;
 KW mitotic checkpoint gene.
 KW Homo sapiens.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1026..3551
 FT /*tag= a
 FT /product= "TTK"
 XX
 PN US2003045491-A1.
 XX 06-MAR-2003.
 XX 21-FEB-2002; 2002US-00081119.
 XX 23-FEB-2001; 2001US-0289813P.
 XX (REIN/) REINHARD C.
 PA (JEFF/) JEFFERSON A B.
 PA (CHAN/) CHAN V W.
 XX Reinhard C, Jefferson AB, Chan VW;
 PI WPI; 2003-455666/43.
 DR P-PSDB; ABU61611.
 XX
 XX Detecting cancer in a subject, by comparing expression levels of tyrosine
 PT threonine kinase polypeptide or polynucleotide in a subject cell and a
 PT normal cell, where an increase in the expression level in the test cell
 PT is indicative of cancer.
 XX
 PS Disclosure; Page 22-25; 79pp; English.
 XX
 CC The invention relates to detecting cancer (other than ovarian cancer) in
 CC a subject, comprising comparing the expression levels of tyrosine
 CC threonine kinase (TTK, a mitotic checkpoint gene) polypeptide or
 CC polynucleotide in a test cell obtained from the subject and in a normal
 CC non-cancer cell, where an increase in the expression level of TTK protein
 CC or nucleic acid in the test cell compared to that in the normal cell,
 CC indicates the presence of cancer other than ovarian cancer. Also included
 CC are reducing growth of a cancerous cell (by contacting a cancerous cell
 CC with an amount of an agent effective to reduce TTK polypeptide activity
 CC in the cell), an assay for identifying a candidate agent that reduces
 CC growth of a cancerous cell (comprising: (i) detecting the activity of a
 CC TTK polypeptide in the presence of a candidate agent; and (ii) comparing
 CC the activity of TTK polypeptide in the presence of a candidate agent
 CC relative to TTK polypeptide activity in the absence of the candidate
 CC agent), identifying an agent that reduces TTK activity (comprising: (i)
 CC contacting a cancerous cell displaying elevated expression of a TTK-
 CC encoding polynucleotide with a candidate agent; and (ii) determining the
 CC effect of the candidate agent on TTK polypeptide activity) and assessing
 CC the prognosis of a cancerous disease other than ovarian cancer in a
 CC subject (comprising: (i) detecting expression of TTK encoding
 CC polynucleotide in a test cancer cell of a subject; and (ii) comparing a
 CC level of expression of TTK-encoding polynucleotide in the test cancer
 CC cell with a level of expression of the polynucleotide in a control non-
 CC cancer cell, where the level of expression of TTK in the test cancer cell
 CC relative to the level of expression in the control non-cancer cell is
 CC indicative of the prognosis of the cancerous disease). The methods are
 CC useful for detecting cancer (other than ovarian cancer) in a subject,
 CC reducing growth of cancerous cells, identifying a candidate agent that
 CC reduces growth of a cancerous cell, identifying an agent that reduces TTK
 CC activity and assessing the prognosis of a cancerous disease other than
 CC ovarian cancer. The methods are also useful for determining the ability
 CC of a subject to respond to a particular therapy e.g. as a basis of
 CC rational therapy. The present sequence encodes human TTK

QY 2101 AAGAACAGAGGTTCCAGAGGTAACAGAAACAGTGGCGAAGCTAAGAGAAAGTCAGAGT 2160
DB |||||
QY 2101 AAGAACAGAGGTTCCAGAGGTAACAGAAACAGTGGCGAAGCTAAGAGAAAGTCAGAGT 2160
DB |||||
QY 2161 GTATTAAACAGAAATCCCTGCTGCATCTTTCAAATCACTGGCAGATTCGGGAGTTAGCCCGAA 2220
DB |||||
QY 2161 GTATTAAACAGAAATCCCTGCTGCATCTTTCAAATCACTGGCAGATTCGGGAGTTAGCCCGAA 2220
DB |||||
QY 2221 AAGTTAATACAGACAGAAATACACCACTTTTGAGCAACCTGCTTTTCAGTTTCAAAC 2280
DB |||||
QY 2221 AAGTTAATACAGACAGAAATACACCACTTTTGAGCAACCTGCTTTTCAGTTTCAAAC 2280
DB |||||
QY 2281 AGTCACCAACCAATATCAACATCTAAATGGTTTGACCCAAATCTATTGTGAAGACCAAA 2340
DB |||||
QY 2281 AGTCACCAACCAATATCAACATCTAAATGGTTTGACCCAAATCTATTGTGAAGACCAAA 2340
DB |||||
QY 2341 GCAGCAATACCTTGGATGATTAATGAGCTGTTTTAGAACTCCAGTTGTAAAGAAATGACT 2400
DB |||||
QY 2341 GCAGCAATACCTTGGATGATTAATGAGCTGTTTTAGAACTCCAGTTGTAAAGAAATGACT 2400
DB |||||
QY 2401 TTCCACCTGCTGTCAGTTGTCACACCTTTATGSCCAACCTGCTGTTTCCAGCAGCAAC 2460
DB |||||
QY 2401 TTCCACCTGCTGTCAGTTGTCACACCTTTATGSCCAACCTGCTGTTTCCAGCAGCAAC 2460
DB |||||
QY 2461 AGCATCAAAATCTTGCCACTCCACTTCAAATTTACAGGTTTTAGCATCTTTCTCAGCAA 2520
DB |||||
QY 2461 AGCATCAAAATCTTGCCACTCCACTTCAAATTTACAGGTTTTAGCATCTTTCTCAGCAA 2520
DB |||||
QY 2521 ATGTAATCCATTTCCGTTAAAGGAGAAATTTATCCATTTAAAGCAGATGAGGAGTGGAG 2580
DB |||||
QY 2521 ATGTAATCCATTTCCGTTAAAGGAGAAATTTATCCATTTAAAGCAGATGAGGAGTGGAG 2580
DB |||||
QY 2581 GTTCAAGCAAGGATTTTCAGGTGTTAAATGAAAGAAACAGATATATGCTTATAAATATG 2640
DB |||||
QY 2581 GTTCAAGCAAGGATTTTCAGGTGTTAAATGAAAGAAACAGATATATGCTTATAAATATG 2640
DB |||||
QY 2641 TGAATTTAAGAGAGAGAGATAACCAACTCTTGATAGTTACCGAAGCAATAGCTTATT 2700
DB |||||
QY 2641 TGAATTTAAGAGAGAGAGATAACCAACTCTTGATAGTTACCGAAGCAATAGCTTATT 2700
DB |||||
QY 2701 TGAATAAATCAACAACACAGTGAAGATCATCCGACTTTATGATTTATGAATCAACGG 2760
DB |||||
QY 2701 TGAATAAATCAACAACACAGTGAAGATCATCCGACTTTATGATTTATGAATCAACGG 2760
DB |||||
QY 2761 ACCAGTACATCATCGTGAATGGAGTGGAAATATTGATCTTAATAGTTGGCTTAAAA 2820
DB |||||
QY 2761 ACCAGTACATCATCGTGAATGGAGTGGAAATATTGATCTTAATAGTTGGCTTAAAA 2820
DB |||||
QY 2821 AGAAATAATCCATTTGATCCATGGAGGAGGAGGAGTTACTGGAAATAATAGTTAGAGCGAG 2880
DB |||||
QY 2821 AGAAATAATCCATTTGATCCATGGAGGAGGAGGAGTTACTGGAAATAATAGTTAGAGCGAG 2880
DB |||||
QY 2881 TTCAACAATCCATCAACATGGCATTTGTTACAGTGTATCTTAAACAGCTAACTTTCTGA 2940
DB |||||
QY 2881 TTCAACAATCCATCAACATGGCATTTGTTACAGTGTATCTTAAACAGCTAACTTTCTGA 2940
DB |||||
QY 2941 TAGTTGATGGAATCTTAAGCTAATTTGATTTGGGATTCGAACCAATTCACCCAGATA 3000
DB |||||
QY 2941 TAGTTGATGGAATCTTAAGCTAATTTGATTTGGGATTCGAACCAATTCACCCAGATA 3000
DB |||||
QY 3001 CAACAAGTGTGTTTAAAGATTTCTCAGGTGGCAGTTAATTTATGCCCACCAAGAGCAA 3060
DB |||||
QY 3001 CAACAAGTGTGTTTAAAGATTTCTCAGGTGGCAGTTAATTTATGCCCACCAAGAGCAA 3060
DB |||||
QY 3061 TCAAGAATATGCTTCTCTCCAGAGAGAAATGGGAAATCTAAGTCAAGATAGAGCCCAAA 3120
DB |||||
QY 3061 TCAAGAATATGCTTCTCTCCAGAGAGAAATGGGAAATCTAAGTCAAGATAGAGCCCAAA 3120
DB |||||
QY 3121 GTGATGTTTGGTCTCTAGGATGATTTTGTACTATATGACTTACGGGAAAAACCACTTTC 3180
DB |||||
QY 3121 GTGATGTTTGGTCTCTAGGATGATTTTGTACTATATGACTTACGGGAAAAACCACTTTC 3180
DB |||||

QY 3181 AGCAGATAATTAATCAGATTTCTAAATTTACATGCGCATAAATGATCCTAATCATGAAATTTG 3240
DB |||||
QY 3181 AGCAGATAATTAATCAGATTTCTAAATTTACATGCGCATAAATGATCCTAATCATGAAATTTG 3240
DB |||||
QY 3241 AATTTCCGATATTTCCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGAGG 3300
DB |||||
QY 3241 AATTTCCGATATTTCCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGAGG 3300
DB |||||
QY 3301 ACCCAAAACAGAGATATCCATTTCTGAGCTCTGGCTCATCCATATGTTTCAAAATTCAAA 3360
DB |||||
QY 3301 ACCCAAAACAGAGATATCCATTTCTGAGCTCTGGCTCATCCATATGTTTCAAAATTCAAA 3360
DB |||||
QY 3361 CTCTATCCAGTTTAAACAAATGCCAAGGGAACCACTGAAGAAATGAATATGTTTCTGGGCC 3420
DB |||||
QY 3361 CTCTATCCAGTTTAAACAAATGCCAAGGGAACCACTGAAGAAATGAATATGTTTCTGGGCC 3420
DB |||||
QY 3421 AACTTGTGTTGCTGANTCTCTAACTCCATTTTGAAGCTGCTAAACCTTTATATCAAC 3480
DB |||||
QY 3421 AACTTGTGTTGCTGANTCTCTAACTCCATTTTGAAGCTGCTAAACCTTTATATCAAC 3480
DB |||||
QY 3481 ACTATAGTGGTGGTCAAAAGTCATAATTTCTTCATCTCCCAAGACTTTTGAAGAAAAAGG 3540
DB |||||
QY 3481 ACTATAGTGGTGGTCAAAAGTCATAATTTCTTCATCTCCCAAGACTTTTGAAGAAAAAGG 3540
DB |||||
QY 3541 GAAAAAATGATTTGCGAGTTATTCGTAATGTCAGATAGGAGGTATATAATATATTGSACT 3600
DB |||||
QY 3541 GAAAAAATGATTTGCGAGTTATTCGTAATGTCAGATAGGAGGTATATAATATATTGSACT 3600
DB |||||
QY 3601 GTTATACTCTTGAATCCCTGFGGAAATCTACATTTGAAGCAACATCACCTCTGAAGTGT 3660
DB |||||
QY 3601 GTTATACTCTTGAATCCCTGFGGAAATCTACATTTGAAGCAACATCACCTCTGAAGTGT 3660
DB |||||
QY 3661 ATCAGCAAAAAAATTCAGTCAGATTTATCTTTAAAGAAAACTGTAAAAATAGCAACCC 3720
DB |||||
QY 3661 ATCAGCAAAAAAATTCAGTCAGATTTATCTTTAAAGAAAACTGTAAAAATAGCAACCC 3720
DB |||||
QY 3721 TTATGGCACTGTATATATTTGATAGACTTTGTTTCTCTGTTTATCTCTTGTGTAATCTAC 3780
DB |||||
QY 3721 TTATGGCACTGTATATATTTGATAGACTTTGTTTCTCTGTTTATCTCTTGTGTAATCTAC 3780
DB |||||
QY 3781 TTGACATCATTTTACTCTTGAATAGTGGGTGAGTACCAAGTATATTCTTAAAAAATTTG 3840
DB |||||
QY 3781 TTGACATCATTTTACTCTTGAATAGTGGGTGAGTACCAAGTATATTCTTAAAAAATTTG 3840
DB |||||
QY 3841 TAAATAAGTGTGCTGCTAAATGA 3866
DB |||||
QY 3841 TAAATAAGTGTGCTGCTAAATGA 3866
DB |||||

RESULT 4
ADD49931
ID ADD49931 standard; cDNA; 3866 BP.
XX
AC ADD49931;
DT
DT 15-JAN-2004 (first entry)
XX
XX Human lung specific tumour antigen L 1479p cDNA.
DE
XX Human; ss; lung cancer antigen; cytostatic; lung cancer; gene therapy;
KW vaccine; T-cell; gene.
XX
OS Homo sapiens.
XX
XX US2003194764-A1.
XX
XX 16-OCT-2003.
XX
XX 04-APR-2002; 2002US-00116712.
XX
XX 05-APR-2001; 2001US-0282289P.
PR
PR 05-OCT-2001; 2001US-0327511P.
XX

QY 1261 TTGAAGGCTTCCCCAGATAAATATGCGCAAAATGAGAGTGTGCTAGAAATCAAGTGA 1320
Db 1261 TTGAAGGCTTCCCCAGATAAATATGCGCAAAATGAGAGTGTGCTAGAAATCAAGTGA 1320
QY 1321 GATTTGCTGAATTAAGAGCTATTCAAGAGCCAGATGATGACGTGACTACTTTCAAATGG 1380
Db 1321 GATTTGCTGAATTAAGAGCTATTCAAGAGCCAGATGATGACGTGACTACTTTCAAATGG 1380
QY 1381 CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCAAAATTTGAATCTGT 1440
Db 1381 CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCAAAATTTGAATCTGT 1440
QY 1441 CACAAGGTAAATGTCAAAAAAGTAAACAACTCTTCAAAAAGCTGTAGAACGTGGAGCAG 1500
Db 1441 CACAAGGTAAATGTCAAAAAAGTAAACAACTCTTCAAAAAGCTGTAGAACGTGGAGCAG 1500
QY 1501 TACCAGTAAATCTCGAAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC 1560
Db 1501 TACCAGTAAATCTCGAAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC 1560
QY 1561 TTTCCAGAGGAGAAAGAGAAATTTATCAGCACTACGGTATTAACCTGCGCAAGATCAT 1620
Db 1561 TTTCCAGAGGAGAAAGAGAAATTTATCAGCACTACGGTATTAACCTGCGCAAGATCAT 1620
QY 1621 TTTCCGGTTCACCTGGGCAATTTACAGAAATAGGAACAACTGTTGATTTCCAGAGGACAGA 1680
Db 1621 TTTCCGGTTCACCTGGGCAATTTACAGAAATAGGAACAACTGTTGATTTCCAGAGGACAGA 1680
QY 1681 CTACTAAAGCAGGTTTTTATATGGAGAGAACATGCCACACAGATGCGAGAAATAGGTT 1740
Db 1681 CTACTAAAGCAGGTTTTTATATGGAGAGAACATGCCACACAGATGCGAGAAATAGGTT 1740
QY 1741 ACCGGAATTCATTGAGACAACTAAACAACTAAACAGTCATGCCATTTGGAAGAGTCC 1800
Db 1741 ACCGGAATTCATTGAGACAACTAAACAACTAAACAGTCATGCCATTTGGAAGAGTCC 1800
QY 1801 CAGTTAACTTTAAATAGCCAGATTGTGATGTAAGACAGATGATCAAGTTGTACCTT 1860
Db 1801 CAGTTAACTTTAAATAGCCAGATTGTGATGTAAGACAGATGATCAAGTTGTACCTT 1860
QY 1861 GTTTTATGAAGACAACTCTAGATCAGAAATGCGAGATTTGGTGTGCTGGAATCTA 1920
Db 1861 GTTTTATGAAGACAACTCTAGATCAGAAATGCGAGATTTGGTGTGCTGGAATCTA 1920
QY 1921 AACCAAGTGAATGATTCGTGTAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
Db 1921 AACCAAGTGAATGATTCGTGTAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
QY 1981 TCAGGAACCTCTGCTCAGATGAAAGAGTTCGAACTTATTAAGTTCATGATTCATTA 2040
Db 1981 TCAGGAACCTCTGCTCAGATGAAAGAGTTCGAACTTATTAAGTTCATGATTCATTA 2040
QY 2041 CCCTGGAAGATAAAACGGAATCAAGTCTTCTAGCTAAATAGAGAACTTAAGAGTATC 2100
Db 2041 CCCTGGAAGATAAAACGGAATCAAGTCTTCTAGCTAAATAGAGAACTTAAGAGTATC 2100
QY 2101 AAGAACAGAGGTTCCAGAGAGTAAACAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAT 2160
Db 2101 AAGAACAGAGGTTCCAGAGAGTAAACAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAT 2160
QY 2161 GTATTAACCAAGATCTCTGTCATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCCGAA 2220
Db 2161 GTATTAACCAAGATCTCTGTCATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCCGAA 2220
QY 2221 AAGTTAATACAGAGCAAAATACCACTTTTGGAGCAACCTGCTTTTTCAGTTTCAAAC 2280
Db 2221 AAGTTAATACAGAGCAAAATACCACTTTTGGAGCAACCTGCTTTTTCAGTTTCAAAC 2280
QY 2281 AGTCACCAACCAATATCAACATCTAAATGTTTGCACCCAAATCTATTTTGAAGACACCA 2340
Db 2281 AGTCACCAACCAATATCAACATCTAAATGTTTGCACCCAAATCTATTTTGAAGACACCA 2340
QY 2341 GCAGCAATACCTTGGATGATTAATGAGCTGTTTTAGAACTCCAGTGTGTAAGAAATGACT 2400

Db 2341 GCAGCAATACCTTGGATGATTAATGAGCTGTTTTAGAACTCCAGTGTGTAAGAAATGACT 2400
QY 2401 TTCCACCTGCTTGTGCTAGTGTCAACCTTATGCGCAACCTGCTGTTTCCAGAGCAAC 2460
Db 2401 TTCCACCTGCTTGTGCTAGTGTCAACCTTATGCGCAACCTGCTGTTTCCAGAGCAAC 2460
QY 2461 AGCATCAATATCTTCCACTCCACTTCAAAATTTACAGGTTTTAGCATCTTCTTCAGCA 2520
Db 2461 AGCATCAATATCTTCCACTCCACTTCAAAATTTACAGGTTTTAGCATCTTCTTCAGCA 2520
QY 2521 ATGAATGCAATTTCCGTTTAAAGGAAGAAATTTATCCATTTAAAGCAGATAGGAAGTGG 2580
Db 2521 ATGAATGCAATTTCCGTTTAAAGGAAGAAATTTATCCATTTAAAGCAGATAGGAAGTGG 2580
QY 2581 GTTCAAGCAAGTATTTTCCAGTGTGTTAAATGAAAGAAACAGATATATGCTATAAATATG 2640
Db 2581 GTTCAAGCAAGTATTTTCCAGTGTGTTAAATGAAAGAAACAGATATATGCTATAAATATG 2640
QY 2641 TGAACCTTAGAAGAGCAGATAACCAACTCTTGTATGTTACCGGAAACGAATAGCTTATT 2700
Db 2641 TGAACCTTAGAAGAGCAGATAACCAACTCTTGTATGTTACCGGAAACGAATAGCTTATT 2700
QY 2701 TGAATAAACTCAACACACAGTGAATAGATCATCCGACTTTATGATTAAGATCAAGG 2760
Db 2701 TGAATAAACTCAACACACAGTGAATAGATCATCCGACTTTATGATTAAGATCAAGG 2760
QY 2761 ACCAGTACATCTAGATGTTATGAGTGTGGAATATTGATTTTAAATAGTGGCTTAA 2820
Db 2761 ACCAGTACATCTAGATGTTATGAGTGTGGAATATTGATTTTAAATAGTGGCTTAA 2820
QY 2821 AGAAAAATTCATTGATCCATGGAAGCAGAGTGTACTGGAAAAATATGTTAGAGCAG 2880
Db 2821 AGAAAAATTCATTGATCCATGGAAGCAGAGTGTACTGGAAAAATATGTTAGAGCAG 2880
QY 2881 TTCACAAATCCATCAACATGGCAATTTTCAAGTGTCTTAAACAGCTTAACTTCTGA 2940
Db 2881 TTCACAAATCCATCAACATGGCAATTTTCAAGTGTCTTAAACAGCTTAACTTCTGA 2940
QY 2941 TAGTTGATGGAATGCTTAAAGCTAAATTTGAGTTTGGGATTCGAAACCAATGCAACAGATA 3000
Db 2941 TAGTTGATGGAATGCTTAAAGCTAAATTTGAGTTTGGGATTCGAAACCAATGCAACAGATA 3000
QY 3001 CAAAGAGTGTGTTAAAGATTTCTCAGGTGCGCAGTAAATTTATATGCCAGCAGAGCA 3060
Db 3001 CAAAGAGTGTGTTAAAGATTTCTCAGGTGCGCAGTAAATTTATATGCCAGCAGAGCA 3060
QY 3061 TCAGAGATATGCTTCTCCAGAGAGATGGAATCTTAAGTCAAGATAAGGCCCAAA 3120
Db 3061 TCAGAGATATGCTTCTCCAGAGAGATGGAATCTTAAGTCAAGATAAGGCCCAAA 3120
QY 3121 GTGATGTTTGGTCTTAGGATGTTATTTGTACTATATGACTTACGGGAAACACCATTTTC 3180
Db 3121 GTGATGTTTGGTCTTAGGATGTTATTTGTACTATATGACTTACGGGAAACACCATTTTC 3180
QY 3181 AGCAGTAATTAATCAGATTTCTAAATTAACATGCTCAATTAATGCTTAAATCAATG 3240
Db 3181 AGCAGTAATTAATCAGATTTCTAAATTAACATGCTCAATTAATGCTTAAATCAATG 3240
QY 3241 AATTTCCGATATTTCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAAAGGG 3300
Db 3241 AATTTCCGATATTTCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAAAGGG 3300
QY 3301 ACCCAAAACAGAGATATCCATTTCTGAGCTCTGCTGCTCATCATATGTTCAAAATCAA 3360
Db 3301 ACCCAAAACAGAGATATCCATTTCTGAGCTCTGCTGCTCATCATATGTTCAAAATCAA 3360
QY 3361 CTCTATCCAGTTTAAACAAATGCGCAAGGAAACCACTGAAGAAATGAAATATGTTCTGGGCC 3420
Db 3361 CTCTATCCAGTTTAAACAAATGCGCAAGGAAACCACTGAAGAAATGAAATATGTTCTGGGCC 3420
QY 3421 AACTTGTGTTGCTGAATTTCTCTTAATCTCCATTTTGAAGCTGCTTAAACCTTATATGAAC 3480

Db 541 CTCCTGGAATCAAGGATCCGCTCCCTCCACTTCCAAAAGTCCCGAGATTACAGGTGTG 600
Qy 601 AGTCACCAATGCTGACCTTAATAATCTTAAGTCAATTTTCTGTCCTATTTCTTCTTAG 660
Db 601 AGTCACCAATGCTGACCTTAATAATCTTAAGTCAATTTTCTGTCCTATTTCTTCTTAG 660
Qy 661 GGTCTCCACAAATCTGCAATAGGGGTGACAAATAATCTTAACCTTCAATGATTCACAAA 720
Db 661 GGTCTCCACAAATCTGCAATAGGGGTGACAAATAATCTTAACCTTCAATGATTCACAAA 720
Qy 721 AGGAGATGAATGATTCATGATTTAGAAAGGGGAAGTAGTAGCCCACTGCACTCCT 780
Db 721 AGGAGATGAATGATTCATGATTTAGAAAGGGGAAGTAGTAGCCCACTGCACTCCT 780
Qy 781 GGATGATGATCCTAAATCCAGATACAGTAAATAATGGGGTATGGGAAGGTAGATACAAA 840
Db 781 GGATGATGATCCTAAATCCAGATACAGTAAATAATGGGGTATGGGAAGGTAGATACAAA 840
Qy 841 TTTGGTTTAAATTAATCTAAATATCTAAATATCTAAACATTTTGGATACANTTTGATGTA 900
Db 841 TTTGGTTTAAATTAATCTAAATATCTAAATATCTAAACATTTTGGATACANTTTGATGTA 900
Qy 901 ATGTAAGACTCTACAGACTTCTAGAAAACAGTTTGGTTCCTATTTTCAATTTCCCCAG 960
Db 901 ATGTAAGACTCTACAGACTTCTAGAAAACAGTTTGGTTCCTATTTTCAATTTCCCCAG 960
Qy 961 TGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGCAGAGAAATGCAAAATGATT 1020
Db 961 TGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGCAGAGAAATGCAAAATGATT 1020
Qy 1021 CCATTAATGCAAAAGTCAGACATTAATAATTAAGTTTAAATTAAGTTCAGTCTTACTGATG 1080
Db 1021 CCATTAATGCAAAAGTCAGACATTAATAATTAAGTTTAAATTAAGTTCAGTCTTACTGATG 1080
Qy 1081 AACTAAGCTTGAATAAAATTTCTGCTGATACACAGATAACTCGGAACTGTAAACAAA 1140
Db 1081 AACTAAGCTTGAATAAAATTTCTGCTGATACACAGATAACTCGGAACTGTAAACAAA 1140
Qy 1141 TTATGATGATGGCAAAACACCCAGAGACTGGTTGAGTTTGGTCTCAAACTAGAGAAA 1200
Db 1141 TTATGATGATGGCAAAACACCCAGAGACTGGTTGAGTTTGGTCTCAAACTAGAGAAA 1200
Qy 1201 ACAGTGTCCCTAAGTGTGCTCTTTTAAATAAATGATTTGCTGTTACAGTCAAGCAA 1260
Db 1201 ACAGTGTCCCTAAGTGTGCTCTTTTAAATAAATGATTTGCTGTTACAGTCAAGCAA 1260
Qy 1261 TTGAAGCGCTTCCCCAGATAAATATGCGCAAAATAGAGTTTGTCTAGAAATTCAGTGA 1320
Db 1261 TTGAAGCGCTTCCCCAGATAAATATGCGCAAAATAGAGTTTGTCTAGAAATTCAGTGA 1320
Qy 1321 GATTTGCTGAATTAAGTATTAACAGAGCCAGATGATGACGTGACTACTTCAATGG 1380
Db 1321 GATTTGCTGAATTAAGTATTAACAGAGCCAGATGATGACGTGACTACTTCAATGG 1380
Qy 1381 CCAGAGCAAACTCGCAAGAAATTTGCTTTTGTATATATCTTTTGCACAAATTTGAATCTG 1440
Db 1381 CCAGAGCAAACTCGCAAGAAATTTGCTTTTGTATATATCTTTTGCACAAATTTGAATCTG 1440
Qy 1441 CACAGCTTAATGTCAAAAGAAAGTAAACACTTCTCAAAAGCTGTAGACGTGGAGCAG 1500
Db 1441 CACAGCTTAATGTCAAAAGAAAGTAAACACTTCTTCAAAAGCTGTAGACGTGGAGCAG 1500
Qy 1501 TACCCTAGAAATGCTCGAAATTTGCCCTGGGAAATTTAAACCTTCCAAAAGAACGACCTGC 1560
Db 1501 TACCCTAGAAATGCTCGAAATTTGCCCTGGGAAATTTAAACCTTCCAAAAGAACGACCTGC 1560
Qy 1561 TTTTCAGAGGAGGAAAAGAGAAATTTATCAGCATCTACGTTATTTAATGCTGCCAAGAAATCAT 1620
Db 1561 TTTTCAGAGGAGGAAAAGAGAAATTTATCAGCATCTACGTTATTTAATGCTGCCAAGAAATCAT 1620
Qy 1621 TTTTCGGGTTTCACTTTGGGCAATTTACAGAAATAGGAACACAGTTTGTGATTTCCAGAGGACAGA 1680
Db 1621 TTTTCGGGTTTCACTTTGGGCAATTTACAGAAATAGGAACACAGTTTGTGATTTCCAGAGGACAGA 1680

Qy 1681 CTACTAAAGCCAGGTTTTTATATGGAGAGAAACATGCCACCAAGATGCAGAAATAGGTT 1740
Db 1681 CTACTAAAGCCAGGTTTTTATATGGAGAGAAACATGCCACCAAGATGCAGAAATAGGTT 1740
Qy 1741 ACCGGAATTCATTTGAGACAAACTTAAACAAACTTAAACAGTCATGCCCAATTTGGAAGAGTCC 1800
Db 1741 ACCGGAATTCATTTGAGACAAACTTAAACAAACTTAAACAGTCATGCCCAATTTGGAAGAGTCC 1800
Qy 1801 CAGTTTAACTTTCTAATATAGCCCAAGTTCTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Db 1801 CAGTTTAACTTTCTAATATAGCCCAAGTTCTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Qy 1861 GTTTTATGAAAAGACAAACCTCTAGATCAGAATGCCAGATTTGGTTGTGCTGCTGATCTA 1920
Db 1861 GTTTTATGAAAAGACAAACCTCTAGATCAGAATGCCAGATTTGGTTGTGCTGCTGATCTA 1920
Qy 1921 AACCAAGTGGAAAATGATTCCTGTGGAATTTAAGAAATTTTAAAGTCTGTTCAAAATAGTCATT 1980
Db 1921 AACCAAGTGGAAAATGATTCCTGTGGAATTTAAGAAATTTTAAAGTCTGTTCAAAATAGTCATT 1980
Qy 1981 TCAAGGAACCTCTGCTGTCAGATGAAAAGAGTCTCTGAACCTTATTATTACTGANTCAATAA 2040
Db 1981 TCAAGGAACCTCTGCTGTCAGATGAAAAGAGTCTCTGAACCTTATTATTACTGANTCAATAA 2040
Qy 2041 CCCTGAGAATAAAAACGGAATCAAGTCTTTAGCTTAAATTTAGAAGAACTAAAGAGTATC 2100
Db 2041 CCCTGAGAATAAAAACGGAATCAAGTCTTTAGCTTAAATTTAGAAGAACTAAAGAGTATC 2100
Qy 2101 AAGAACAGAGGTTCCAGAGAGTAAACAGAAACAGTGCAGCTTAAGAGAAAGTCAGAGT 2160
Db 2101 AAGAACAGAGGTTCCAGAGAGTAAACAGAAACAGTGCAGCTTAAGAGAAAGTCAGAGT 2160
Qy 2161 GTATTAAACAGAAATCCTGCTGCATCTTTCAATCTACTGCGAGATTTCCGGAGTTAGCCCGAA 2220
Db 2161 GTATTAAACAGAAATCCTGCTGCATCTTTCAATCTACTGCGAGATTTCCGGAGTTAGCCCGAA 2220
Qy 2221 AAGTTAATACAGAGAGAAACATACCCTTTTGAAGAACCTGCTCTTTTCAAGTTCAAAAC 2280
Db 2221 AAGTTAATACAGAGAGAAACATACCCTTTTGAAGAACCTGCTCTTTTCAAGTTTCAAAAC 2280
Qy 2281 AGTCACCAACCAATCAACACTTAAATGGTTTGAACCAAAATCTATTGTGAAGACACAA 2340
Db 2281 AGTCACCAACCAATCAACACTTAAATGGTTTGAACCAAAATCTATTGTGAAGACACAA 2340
Qy 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTTAGAACTCCAGTTGTAAGAAATGACT 2400
Db 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTTAGAACTCCAGTTGTAAGAAATGACT 2400
Qy 2401 TTCCACCTGCTTGTGTCAGTTGTCACACCTTATGGCCCACTGCTGTTTCCAGCAGCAAC 2460
Db 2401 TTCCACCTGCTGCTGTCAGTTGTCACACCTTATGGCCCACTGCTGTTTCCAGCAGCAAC 2460
Qy 2461 AGCATCAAAATCTTGGCACTCCACTTCAAAATTTACAGTTTACAGTCTTCTTCAGCAA 2520
Db 2461 AGCATCAAAATCTTGGCACTCCACTTCAAAATTTACAGTTTACAGTCTTCTTCAGCAA 2520
Qy 2521 ATGATGCAATTCGGTTTAAAGGAGAAATTTATTCATATTTAAAGCAGATAGGAGTGGAG 2580
Db 2521 ATGATGCAATTCGGTTTAAAGGAGAAATTTATTCATATTTAAAGCAGATAGGAGTGGAG 2580
Qy 2581 GTTCAAGCAAGGATTTTTCAGGTGTTAAATGAAAAGAAACAGATATATGCTATATAAATATG 2640
Db 2581 GTTCAAGCAAGGATTTTTCAGGTGTTAAATGAAAAGAAACAGATATATGCTATATAAATATG 2640
Qy 2641 TGAATTTAGAGAGAGAGATCAACAACTCTTGTAGTTACCGGAACGAATAGCTTTATT 2700
Db 2641 TGAATTTAGAGAGAGAGATCAACAACTCTTGTAGTTACCGGAACGAATAGCTTTATT 2700
Qy 2701 TGAATAAACTACAAACACAGATGATAGATCATCCGACTTTTATGATTTATGAAATCACGG 2760
Db 2701 TGAATAAACTACAAACACAGATGATAGATCATCCGACTTTTATGATTTATGAAATCACGG 2760

QY 2761 ACCAGTACATCTACATCGTAAATGGAGTGTGGAATATTTGATCTTAATAGTGGCTTAAAA 2820
 Db 2761 ACCAGTACATCTACATCGTAAATGGAGTGTGGAATATTTGATCTTAATAGTGGCTTAAAA 2820
 QY 2821 AGAAAAATCCATTCATCCATGGACGCAAGAGTTACTGGAAAAATATCTTAGAGGCAG 2880
 Db 2821 AGAAAAATCCATTCATCCATGGACGCAAGAGTTACTGGAAAAATATCTTAGAGGCAG 2880
 QY 2881 TTCACAATCCATTCATCCATGGACGCAAGAGTTACTGGAAAAATATCTTAGAGGCAG 2940
 Db 2881 TTCACAATCCATTCATCCATGGACGCAAGAGTTACTGGAAAAATATCTTAGAGGCAG 2940
 QY 2941 TAGTTGATGAATGCTTAAGCTAATTCATTTTGGGATTCGAAACCAATGCACACAGATA 3000
 Db 2941 TAGTTGATGAATGCTTAAGCTAATTCATTTTGGGATTCGAAACCAATGCACACAGATA 3000
 QY 3001 CAACAAGTGTGTTAAAGATTCACAGTTTGGACAGTTAATATATGCCACCAAGAGCAA 3060
 Db 3001 CAACAAGTGTGTTAAAGATTCACAGTTTGGACAGTTAATATATGCCACCAAGAGCAA 3060
 QY 3061 TCAAGATATGCTTCTCCAGAGAGAAATCTAAGTCAAGATAGCCCAAAA 3120
 Db 3061 TCAAGATATGCTTCTCCAGAGAGAAATCTAAGTCAAGATAGCCCAAAA 3120
 QY 3121 GTGATGTTTGTCTTGGATGATTTTGTACTATATGACTTACGGGAAACACCATTTTC 3180
 Db 3121 GTGATGTTTGTCTTGGATGATTTTGTACTATATGACTTACGGGAAACACCATTTTC 3180
 QY 3181 AGCAGATAATTAATCAGATTCTAAATTAATCATGCCATTAATGATCTTAATCATGAATG 3240
 Db 3181 AGCAGATAATTAATCAGATTCTAAATTAATCATGCCATTAATGATCTTAATCATGAATG 3240
 QY 3241 AATTTCCGATATTCAGAGAGAGATCTTCAAGATGTTAAAGTGTGTTTAAAAAGGG 3300
 Db 3241 AATTTCCGATATTCAGAGAGAGATCTTCAAGATGTTAAAGTGTGTTTAAAAAGGG 3300
 QY 3301 ACCCAACAGAGATATCCATTCCTGAGCTCTGCTCATCCATATGTTTCAAAATCAA 3360
 Db 3301 ACCCAACAGAGATATCCATTCCTGAGCTCTGCTCATCCATATGTTTCAAAATCAA 3360
 QY 3361 CTCATCAGTTAACCAATGCCAAGGAGACCATGAGAAATGAATATGTTCTGGGCC 3420
 Db 3361 CTCATCAGTTAACCAATGCCAAGGAGACCATGAGAAATGAATATGTTCTGGGCC 3420
 QY 3421 AACTTGTGCTGAATTCCTTCACTCACTTCTTGAAGCTGCTTAAACCTTTATATGAAC 3480
 Db 3421 AACTTGTGCTGAATTCCTTCACTCACTTCTTGAAGCTGCTTAAACCTTTATATGAAC 3480
 QY 3481 ACTATAGTGGTGAAGTCAATATCTTCACTCCTCAAGCTCTTGAAGCTTGAAGCT 3540
 Db 3481 ACTATAGTGGTGAAGTCAATATCTTCACTCCTCAAGCTCTTGAAGCTTGAAGCT 3540
 QY 3541 GAAAAAATGATTTGCGATTTATCGTAAATGCTAGATAGGAGGTATATAATATTTGGACT 3600
 Db 3541 GAAAAAATGATTTGCGATTTATCGTAAATGCTAGATAGGAGGTATATAATATTTGGACT 3600
 QY 3601 GTTATCTCTTGAATCCCTGTGGAATCTACATTTGAAGCAACATCACTCTGAAGTGT 3660
 Db 3601 GTTATCTCTTGAATCCCTGTGGAATCTACATTTGAAGCAACATCACTCTGAAGTGT 3660
 QY 3661 ATCAGCAAAAAAATTCAGTGAATATCTTAAAGAAACTGTAAATAGCAACAC 3720
 Db 3661 ATCAGCAAAAAAATTCAGTGAATATCTTAAAGAAACTGTAAATAGCAACAC 3720
 QY 3721 TTATGGCACTGTATATTTGAGACTGTTTCTCTGTTTATGCTCTTGTGTAATCTAC 3780
 Db 3721 TTATGGCACTGTATATTTGAGACTGTTTCTCTGTTTATGCTCTTGTGTAATCTAC 3780
 QY 3781 TTGACATCAATTTTACTCTTGAATAGTGGGTGAGTATATTTCTTAAAAAATCTTTG 3840
 Db 3781 TTGACATCAATTTTACTCTTGAATAGTGGGTGAGTATATTTCTTAAAAAATCTTTG 3840
 QY 3841 TAAATAAGATTTTCTGGCTAAATGA 3866

Db 3841 TAAATAAGATTTTCTGGCTAAATGA 3866
 RESULT 6
 ADE38368
 ID ADE38368 standard; DNA; 3866 BP.
 XX ADE38368;
 XX 29-JAN-2004 (first entry)
 DE Human protein 1682 gene sequence.
 XX tumorigenic disorder; angiogenic disorder; aberrant gene expression;
 KW aberrant protein activity; cytostatic; antithyroid; antidiabetic;
 KW ophthalmological; cancer; breast cancer; colon cancer; lung cancer;
 KW prostatic cancer; Grave's disease; diabetic retinopathy; gene; ds;
 XX protein 1682.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 978..3551
 FT /*tag= a
 ET /product= "Human protein 1682"
 XX
 PN WO2003065006-A2.
 XX
 PD 07-AUG-2003.
 XX
 PF 30-JAN-2003; 2003WO-US002588.
 XX
 PR 31-JAN-2002; 2002US-0353600P.
 PR 15-MAR-2002; 2002US-0364517P.
 PR 09-APR-2002; 2002US-0371075P.
 PR 10-APR-2002; 2002US-0371507P.
 PR 16-APR-2002; 2002US-0372984P.
 PR 19-APR-2002; 2002US-0374194P.
 PR 24-MAY-2002; 2002US-0382995P.
 PR 31-MAY-2002; 2002US-0385023P.
 PR 14-JUN-2002; 2002US-0388853P.
 PR 17-JUN-2002; 2002US-0389395P.
 PR 25-JUN-2002; 2002US-0391324P.
 PR 15-JUL-2002; 2002US-0395944P.
 PR 22-JUL-2002; 2002US-0397726P.
 PR 13-AUG-2002; 2002US-0403046P.
 PR 22-AUG-2002; 2002US-0403155P.
 PR 27-AUG-2002; 2002US-0408361P.
 PR 25-OCT-2002; 2002US-0421195P.
 PR 12-NOV-2002; 2002US-0425456P.
 PR 19-NOV-2002; 2002US-0427626P.
 PR 10-DEC-2002; 2002US-0432122P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Hunter JJ, Macbeth KJ, Tsai F, Lesoon A, Lightcap ES;
 PI Williamson MW, Rudolph-Owen LA;
 XX
 XX WPI: 2003-646176/61.
 DR P-PSDB; ADE38369.
 DR
 XX
 PT Treating subject having tumorigenic disorder or angiogenic disorder
 PT caused by aberrant polypeptide e.g., N-formylpeptide receptor or nucleic
 PT acid, by administering a modulator.
 XX
 PS Disclosure; SEQ ID NO 29; 454pp; English.
 CC
 CC This invention relates to a novel method of treating a human subject
 CC having a tumorigenic disorder or angiogenic disorder, caused by aberrant
 CC gene expression or activity of an isolated protein, by administering a
 CC modulator. The modulator may have cytostatic, antithyroid, antidiabetic
 CC or ophthalmological activity. The method is useful for treating a subject

CC having a tumorigenic or angiogenic disorder, in particular for treating
CC cancer (for example breast cancer, colon cancer, lung cancer or prostatic
CC cancer) and, for example, Grave's disease and diabetic retinopathy. The
CC present sequence is a DNA sequence which encodes the novel isolated human
CC protein 1682 of the invention.

SQ Sequence 3866 BP; 1268 A; 731 C; 752 G; 1115 T; 0 U; 0 Other;
 Query Match 100.0%; Score 3866; DB 9; Length 3866;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3866; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGAATTCCTTTTTTTTTTTTTTTTTCAGATGGAGTTTCACTCTGTGTGCCAGGCTGGAGTG 60
 DB 1 GGAATTCCTTTTTTTTTTTTTTTTTCAGATGGAGTTTCACTCTGTGTGCCAGGCTGGAGTG 60
 QY 61 CAATGGCACAAATCTCAGCTTACTGCAACCTCCGCTCCCGGTTTCAAGCGATTCTCCTGC 120
 DB 61 CAATGGCACAAATCTCAGCTTACTGCAACCTCCGCTCCCGGTTTCAAGCGATTCTCCTGC 120
 QY 121 CTCAGCCTCTCAAGTAGCTGGATTACAGGCATGTGCCACACCCCTGGCTAACTAATTT 180
 DB 121 CTCAGCCTCTCAAGTAGCTGGATTACAGGCATGTGCCACACCCCTGGCTAACTAATTT 180
 QY 181 CTTTCTCTATTAGTAGAGATGGGGTTTTCACCATTTGGTCAGGCTGGTCTTGAACCTCGT 240
 DB 181 CTTTCTCTATTAGTAGAGATGGGGTTTTCACCATTTGGTCAGGCTGGTCTTGAACCTCGT 240
 QY 241 ACCTCAGGTGATCCACTTGCCTTGGCCCTCCAAAGTCTAGGATTACAGCGTGAACCTG 300
 DB 241 ACCTCAGGTGATCCACTTGCCTTGGCCCTCCAAAGTCTAGGATTACAGCGTGAACCTG 300
 QY 301 TGCCTGGCTGATCTCTTTTTTTGTTTGGATTTTTTGAAACAGGGTCTCCCTTGGTCGCC 360
 DB 301 TGCCTGGCTGATCTCTTTTTTTGTTTGGATTTTTTGAAACAGGGTCTCCCTTGGTCGCC 360
 QY 361 AGCGTGAGTGCAGTGGTGCGATCTTGGCTCCACTATAACCTCCACCTCTGGTTTCAAGT 420
 DB 361 AGCGTGAGTGCAGTGGTGCGATCTTGGCTCCACTATAACCTCCACCTCTGGTTTCAAGT 420
 QY 421 GATCCTCCCACTTTAGCCTCTCAGTAGCTGTGATTAACAGCGGTGCACCAACACCCGG 480
 DB 421 GATCCTCCCACTTTAGCCTCTCAGTAGCTGTGATTAACAGCGGTGCACCAACACCCGG 480
 QY 481 CTAATTTTGTGATTTTTATTAGACAGGGTTTCCACCATGTGGCCAGGCTGTCTCAAA 540
 DB 481 CTAATTTTGTGATTTTTATTAGACAGGGTTTCCACCATGTGGCCAGGCTGTCTCTCAAA 540
 QY 541 CTCCTGGACTCAAGGGATCCGCTGCCCTCCACTTCCAAAAGTCCGAGATTACAGGTGTG 600
 DB 541 CTCCTGGACTCAAGGGATCCGCTGCCCTCCACTTCCAAAAGTCCGAGATTACAGGTGTG 600
 QY 601 AGTCACCATGCGTCGACCTTAATCTTAACTTAAGTCAATTTTTTCTGGTCAATTTCTCTAG 660
 DB 601 AGTCACCATGCGTCGACCTTAATCTTAACTTAAGTCAATTTTTTCTGGTCAATTTCTCTAG 660
 QY 661 GGTCCCTCACAACAATCTGCATTAAGCGGTACAAATACTCTAACTCATGATTCACAAA 720
 DB 661 GGTCCCTCACAACAATCTGCATTAAGCGGTACAAATACTCTAACTCATGATTCACAAA 720
 QY 721 AGGAAGATGAAGTGATTCATGATTTAGAAAGGGAGTAGTAGCCCATCTGCACACTCCT 780
 DB 721 AGGAAGATGAAGTGATTCATGATTTAGAAAGGGAGTAGTAGCCCATCTGCACACTCCT 780
 QY 781 GGATGATGATCTTAAATCCAGATACAGTAAAAATGGGGTATGGGAAGGTAGAATACAAA 840
 DB 781 GGATGATGATCTTAAATCCAGATACAGTAAAAATGGGGTATGGGAAGGTAGAATACAAA 840
 QY 841 TTTTGGTTTAAATTAATTTATCTTAATATCTAAAAACATTTTTGGATACATTTGTGTGATGA 900
 DB 841 TTTTGGTTTAAATTAATTTATCTTAATATCTAAAAACATTTTTGGATACATTTGTGTGATGA 900
 QY 901 ATGTAAGACTGTACAGACTTCTCTAGAAAAACAGTTTGGGTTCCATCTTTTTCAATTTCCCG 960

Db 1981 TCAGGAACCTCTGGTGCAGATGAAAGAGTCTTGAACCTATTATTACTGATTCATTA 2040
Qy
2041 CCCTGAAGAAATAAAACGGAATCAAGCTCTTCTAGCTAAATTTAGAGAAACCTAAGAGATTC 2100
Db
2041 CCCTGAAGAAATAAAACGGAATCAAGCTCTTCTAGCTAAATTTAGAGAAACCTAAGAGATTC 2100
Qy
2101 AAGAACACAGAGGTTCCAGAGAGTAAACAGAGAAACAGTGGCAAGCTAAGAGAGAAAGTCAGAGT 2160
Db
2101 AAGAACACAGAGGTTCCAGAGAGTAAACAGAGAAACAGTGGCAAGCTAAGAGAGAAAGTCAGAGT 2160
Qy
2161 GTATTAAACAGAAATCTCTGCTGCATCTTTCAAAATCACTGGCAGATCCGGAGTTAGCCCGAA 2220
Db
2161 GTATTAAACAGAAATCTCTGCTGCATCTTTCAAAATCACTGGCAGATCCGGAGTTAGCCCGAA 2220
Qy
2221 AAGTTAATACAGAGCAGAAACATACCACTTTTCAGCAACCTGCTTTTCAGTTTCAAAAC 2280
Db
2221 AAGTTAATACAGAGCAGAAACATACCACTTTTCAGCAACCTGCTTTTCAGTTTCAAAAC 2280
Qy
2281 AGTCACCAACAATATCAACATCTAAATGGTTTGAACCCAAATCTATTGTGTAAGACACCAA 2340
Db
2281 AGTCACCAACAATATCAACATCTAAATGGTTTGAACCCAAATCTATTGTGTAAGACACCAA 2340
Qy
2341 GCAGCAATACCTGGATGATTAACATGAGCTGTTTGAACCTCCAGTTGTAAGAAATGACT 2400
Db
2341 GCAGCAATACCTGGATGATTAACATGAGCTGTTTGAACCTCCAGTTGTAAGAAATGACT 2400
Qy
2401 TTCCACCTGCTGTGACGTGTCACACCTTATGGCCAACTGCTGTTTCCAGCAGCAAC 2460
Db
2401 TTCCACCTGCTGTGACGTGTCACACCTTATGGCCAACTGCTGTTTCCAGCAGCAAC 2460
Qy
2461 AGCATCAAAATCTTGCACCTCCACTTCAAAATTTACAGGTTTTCAGATCTTCTTCAGCAA 2520
Db
2461 AGCATCAAAATCTTGCACCTCCACTTCAAAATTTACAGGTTTTCAGATCTTCTTCAGCAA 2520
Qy
2521 ATGAATGCATTTCCGTTTAAAGGAAGATTTATTCATATTAAGCAGATAGGAGTGGAG 2580
Db
2521 ATGAATGCATTTCCGTTTAAAGGAAGATTTATTCATATTAAGCAGATAGGAGTGGAG 2580
Qy
2581 GTTCAAGCAAGGTTATTCAGGTGTTAAATGAAAGAAACAGATATATGCTATTAATAATATG 2640
Db
2581 GTTCAAGCAAGGTTATTCAGGTGTTAAATGAAAGAAACAGATATATGCTATTAATAATATG 2640
Qy
2641 TGAACTTAGAGAGAGAGTAAACCAACTCTTGATAGTTACCGGACGAATAGCTTATT 2700
Db
2641 TGAACTTAGAGAGAGAGTAAACCAACTCTTGATAGTTACCGGACGAATAGCTTATT 2700
Qy
2701 TGAATAAATCAACAACACAGTGAATAGATCATCCGACTTTATGATTAATGAATCAGCG 2760
Db
2701 TGAATAAATCAACAACACAGTGAATAGATCATCCGACTTTATGATTAATGAATCAGCG 2760
Qy
2761 ACCAGTACATCTAGATGTAAGAGTGGAAATATGATCTTAATAGTGGCTTAAAA 2820
Db
2761 ACCAGTACATCTAGATGTAAGAGTGGAAATATGATCTTAATAGTGGCTTAAAA 2820
Qy
2821 AGAAAAATCCATTGATCCATGGGAAACGCAAGAGTTACTGGAAAAATATGTTAGAGCAG 2880
Db
2821 AGAAAAATCCATTGATCCATGGGAAACGCAAGAGTTACTGGAAAAATATGTTAGAGCAG 2880
Qy
2881 TTCACACAATCCATCAACATGGCAATGTTCAAGTGTCTTAAACCGAGCTAACTTCTGA 2940
Db
2881 TTCACACAATCCATCAACATGGCAATGTTCAAGTGTCTTAAACCGAGCTAACTTCTGA 2940
Qy
2941 TAGTTGATGGAATGCTAAGGCTAATGATTTTGGGATTCGAAACCAATGCAACCCAGATA 3000
Db
2941 TAGTTGATGGAATGCTAAGGCTAATGATTTTGGGATTCGAAACCAATGCAACCCAGATA 3000
Qy
3001 CAACAAGTGTGTTAAAGATTTCTCAGGTTGGCAGAGTTAATATATATGCCACCAAGACAA 3060
Db
3001 CAACAAGTGTGTTAAAGATTTCTCAGGTTGGCAGAGTTAATATATATGCCACCAAGACAA 3060
Qy
3061 TCAAGAGATATGCTTCTCCAGAGAGAAATGGGAAATCTAAGTCAAGATAGCCCAAAA 3120
Db
3061 TCAAGAGATATGCTTCTCCAGAGAGAAATGGGAAATCTAAGTCAAGATAGCCCAAAA 3120

Qy 3121 GTGATGTTTGGTCTTAGGATGTTTTTGTACTATATGACTTACGGGAAAAACACCATTTTC 3180
Db 3121 GTGATGTTTGGTCTTAGGATGTTTTTGTACTATATGACTTACGGGAAAAACACCATTTTC 3180
Qy 3181 AGCAGATAAATTAATCAGATTTTCTAAATFACATGCCATAAATGATCCTTAATCATGAAATG 3240
Db 3181 AGCAGATAAATTAATCAGATTTTCTAAATFACATGCCATAAATGATCCTTAATCATGAAATG 3240
Qy 3241 AATTTCCGATATCCAGAGAAAGATCTTCAAGATGCTGTTAAAGTGTGTTTAAAAAGGG 3300
Db 3241 AATTTCCGATATCCAGAGAAAGATCTTCAAGATGCTGTTAAAGTGTGTTTAAAAAGGG 3300
Qy 3301 ACCCAAAAACAGAGATATCCATTCCTGAGCTCCTGGCTCATCCATATGTTCAAAATTCAAA 3360
Db 3301 ACCCAAAAACAGAGATATCCATTCCTGAGCTCCTGGCTCATCCATATGTTCAAAATTCAAA 3360
Qy 3361 CTCATCCAGTTAAACAAATGGCAAGGAAACCACTGAAGAAATGAAATATGTTCTGGGCC 3420
Db 3361 CTCATCCAGTTAAACAAATGGCAAGGAAACCACTGAAGAAATGAAATATGTTCTGGGCC 3420
Qy 3421 AACTTTGGTGTGTAATCTCCTAACTCCATTTTGAAGAGCTGCTAAAACTTTATATGAAC 3480
Db 3421 AACTTTGGTGTGTAATCTCCTAACTCCATTTTGAAGAGCTGCTAAAACTTTATATGAAC 3480
Qy 3481 ACTATAGTGGTGTGAAAGTCATAATCTTCACTCCCAAGACTTTTGAAAAAAGGG 3540
Db 3481 ACTATAGTGGTGTGAAAGTCATAATCTTCACTCCCAAGACTTTTGAAAAAAGGG 3540
Qy 3541 GAARAAAATGATTTCCAGTTAATTCGTAATGTCAGATAGGAGTATAAATATATTGGACT 3600
Db 3541 GAARAAAATGATTTCCAGTTAATTCGTAATGTCAGATAGGAGTATAAATATATTGGACT 3600
Qy 3601 GTTATACTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAACATCACTCTGAAGTGT 3660
Db 3601 GTTATACTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAACATCACTCTGAAGTGT 3660
Qy 3661 ATCAGCAAAAAAATTCAGTGAGATTTCTTTAAAGAAATCTGTAAGAAATAGCAACAC 3720
Db 3661 ATCAGCAAAAAAATTCAGTGAGATTTCTTTAAAGAAATCTGTAAGAAATAGCAACAC 3720
Qy 3721 TTATGGCACTGTATATATTGTAGACTTGTCTCTGTTTTATGCTCTTGTGTAATCTAC 3780
Db 3721 TTATGGCACTGTATATATTGTAGACTTGTCTCTGTTTTATGCTCTTGTGTAATCTAC 3780
Qy 3781 TTGACATCAATTTTACTCTTGGAAATAGTGGTGGATAGCAAGTATATTTCTAAAAAATTTG 3840
Db 3781 TTGACATCAATTTTACTCTTGGAAATAGTGGTGGATAGCAAGTATATTTCTAAAAAATTTG 3840
Qy 3841 TAAATAAAGTTTTTGGGCTAAAAATGA 3866
Db 3841 TAAATAAAGTTTTTGGGCTAAAAATGA 3866

RESULT 7

ADE31644

ID ADE31644 standard; DNA; 3866 BP.

XX

AC ADE31644;

XX

DT 29-JAN-2004 (first entry)

XX

DE Human 1682 gene #SEQ ID 1.

XX

KW Antiarteriosclerotic; cardiant; vasotropic; antiinflammatory;

KW thrombolytic; antiarrhythmic; antianginal; hypotensive; gene therapy;

KW cardiovascular; disorder; ischaemia; aortic bending;

KW vascular heart disease; endocarditis; atrial fibrillation; heart failure;

XX angina; cardiomyopathy; cardiac death; gene; ds.

OS Homo sapiens.

XX

PN WO2003065984-A2.

QY 1705 GAGAGACATGCCACCAAGATGCAGAAATAGGTTACCGGAATTCATTGAGCAAACTA 1764
DB 811 GAGAGACATGCCACCAAGATGCAGAAATAGGTTACCGGAATTCATTGAGCAAACTA 870
QY 1765 ACAAACCTAAACAGTCATGCCATTTGGAAAGAGTCCAGATTAAACCTTCTAAATAGCCGAG 1824
DB 871 ACAAACCTAAACAGTCATGCCATTTGGAAAGAGTCCAGATTAAACCTTCTAAATAGCCGAG 930
QY 1825 ATTGTGATGGAAGACAGATGATTCAGTTGTGATCTTGTGTTTATGAAAGACAAACCTCTA 1884
DB 931 ATTGTGATGGAAGACAGATGATTCAGTTGTGATCTTGTGTTTATGAAAGACAAACCTCTA 990
QY 1885 GATCAGAAATCGGAGATTGGTTGTGCTCGATCTAAACCAAGTGGAAATGATTCCTGTG 1944
DB 991 GATCAGAAATCGGAGATTGGTTGTGCTCGATCTAAACCAAGTGGAAATGATTCCTGTG 1050
QY 1945 AATTAAGAAATTTAAAGCTGTCTCAAAATAGTCATTTCAAGAACCTCTGCTCAGATG 2004
DB 1051 AATTAAGAAATTTAAAGCTGTCTCAAAATAGTCATTTCAAGAACCTCTGCTCAGATG 1110
QY 2005 AAAAGGTTCTGAACCTTATTACTGATTCATTAACCTGAAGATTAAGCGGATCAA 2064
DB 1111 AAAAGGTTCTGAACCTTATTACTGATTCATTAACCTGAAGATTAAGCGGATCAA 1170
QY 2065 GTCTTCTAGCTAAATTAGAGAAACTAAAGAGTATCAAGAACCCAGAGGTTCCAGAGAGTA 2124
DB 1171 GTCTTCTAGCTAAATTAGAGAAACTAAAGAGTATCAAGAACCCAGAGGTTCCAGAGAGTA 1230
QY 2125 ACAGAAACAGTGCAGACTTAAGAAAGTCAGAGTGTTTAAACCAGATCTCTGCTGCAT 2184
DB 1231 ACCAGAAACAGTGCAGACTTAAGAAAGTCAGAGTGTTTAAACCAGATCTCTGCTGCAT 1290
QY 2185 CTTCAAAATCACTGCGAGATTCGGAGTTAGCCGAAAAGTTAATACAGACAGAAACATA 2244
DB 1291 CTTCAAAATCACTGCGAGATTCGGAGTTAGCCGAAAAGTTAATACAGACAGAAACATA 1350
QY 2245 CCACCTTTGAGCAACCTGTCTTTTCAAGTTCAAAACAGTCACCAACCAATATCAACATCTA 2304
DB 1351 CCACCTTTGAGCAACCTGTCTTTTCAAGTTCAAAACAGTCACCAACCAATATCAACATCTA 1410
QY 2305 AATGGTTGACCCAAATCTATTGTGAAGACACCAAGCAGCAATACCTTGGATGATTACA 2364
DB 1411 AATGGTTGACCCAAATCTATTGTGAAGACACCAAGCAGCAATACCTTGGATGATTACA 1470
QY 2365 TGAGCTGTTTAGAACCTCCAGTTGTAAGAAATGACTTTCCACCTGCTTGTCACTGTCAA 2424
DB 1471 TGAGCTGTTTAGAACCTCCAGTTGTAAGAAATGACTTTCCACCTGCTTGTCACTGTCAA 1530
QY 2425 CACCTTATGGCCAACTGCTGTTTCCAGAGCAAGCAGATCAAAATCTTGCACCTCCAC 2484
DB 1531 CACCTTATGGCCAACTGCTGTTTCCAGAGCAAGCAGATCAAAATCTTGCACCTCCAC 1590
QY 2485 TTCAAAATTTACAGGTTTACATCTCTTCCAGCAAAATGAATGCAATTCGGTTAAAGGAA 2544
DB 1591 TTCAAAATTTACAGGTTTACATCTCTTCCAGCAAAATGAATGCAATTCGGTTAAAGGAA 1650
QY 2545 GAATTTATTCOATTTAAAGCAGATAGGAAGTGGAGTTCAAGCAAGGATTTTCAGGTGT 2604
DB 1651 GAATTTATTCOATTTAAAGCAGATAGGAAGTGGAGTTCAAGCAAGGATTTTCAGGTGT 1710
QY 2605 TAAATGAAGAAACAGATATGCTATATAAATGTGAATTTAGAACTTAGAAGACAGATAACC 2664
DB 1711 TAAATGAAGAAACAGATATGCTATATAAATGTGAATTTAGAACTTAGAAGACAGATAACC 1770
QY 2665 AAACTTTGATAGTTACCGGAACGAAATAGCTTTTGAATTAACCTCAACAACACAGTG 2724
DB 1771 AAACTTTGATAGTTACCGGAACGAAATAGCTTTTGAATTAACCTCAACAACACAGTG 1830
QY 2725 ATAGATCATCCGACTTTATGATTTGAATTCAGGACGACAGTACATCTACATGTTATGG 2784
DB 1831 ATAGATCATCCGACTTTATGATTTGAATTCAGGACGACAGTACATCTACATGTTATGG 1890
QY 2785 AGTGTGGAATATTGATCTTAATAGTTGGCTTAAAGAAAGAAAAATCCATTGATCCATGGG 2844

DB 1891 AGTGTGGAATATTGATCTTTAATAGTTGGCTTAAAGAAAGAAAAATCCATTGATCCATGGG 1950
QY 2845 AACCAAGAGATTACTGGAAAAATATGTTAGAGGAGATTTCACACAATCATCAACATGCA 2904
DB 1951 AACCAAGAGATTACTGGAAAAATATGTTAGAGGAGATTTCACACAATCATCAACATGCA 2010
QY 2905 TTGTTCCAGTGATCTTAAACCAAGTAACTTTCTGATAGTTGATGGAATGCTTAAGCTAA 2964
DB 2011 TTGTTCCAGTGATCTTAAACCAAGTAACTTTCTGATAGTTGATGGAATGCTTAAGCTAA 2070
QY 2965 TTGATTTTGGGATTCAAAACCAAGTAACTTCACAGATACAAAGTGTGTTTAAAGATTCTC 3024
DB 2071 TTGATTTTGGGATTCAAAACCAAGTAACTTCACAGATACAAAGTGTGTTTAAAGATTCTC 2130
QY 3025 AGTTGGGACAGTTAATTAATATATGCCACAGAGCAATCAAGATATGTTCTTCCACAG 3084
DB 2131 AGTTGGGACAGTTAATTAATATATGCCACAGAGCAATCAAGATATGTTCTTCCACAG 2190
QY 3085 AGAATGGGAATCTAAAGTCAAGATAAGCCCAAGTGAATGTTTGGTCTCTTAGGATGTA 3144
DB 2191 AGAATGGGAATCTAAAGTCAAGATAAGCCCAAGTGAATGTTTGGTCTCTTAGGATGTA 2250
QY 3145 TTTTGTACTATATGACTTACGGGAAAAACCAATTTACAGAGATAATTAATCAGATTCTCA 3204
DB 2251 TTTTGTACTATATGACTTACGGGAAAAACCAATTTACAGAGATAATTAATCAGATTCTCA 2310
QY 3205 RAATACATGCCATAATGATCTCTAATCATGAATTTCCCGATATTTCCAGAGAAAG 3264
DB 2311 RAATACATGCCATAATGATCTCTAATCATGAATTTCCCGATATTTCCAGAGAAAG 2370
QY 3265 ATCTTCAAGATGTGTTAAAGTGTGTTTAAAGAGGACCCAAAACAGAGGATATCCATTC 3324
DB 2371 ATCTTCAAGATGTGTTAAAGTGTGTTTAAAGAGGACCCAAAACAGAGGATATCCATTC 2430
QY 3325 CTGAGCTCTGCTCATCCATATGTTCAAAATCAAACTCATCCAGTAAACCAATGGGCA 3384
DB 2431 CTGAGCTCTGCTCATCCATATGTTCAAAATCAAACTCATCCAGTAAACCAATGGGCA 2490
QY 3385 AGGGAACCACTCAAGAAATGAAATATGTTTGGGCAAACTTGTGGTCTGAAATTCCTCA 3444
DB 2491 AGGGAACCACTCAAGAAATGAAATATGTTTGGGCAAACTTGTGGTCTGAAATTCCTCA 2550
QY 3445 ACTCCATTTTGAAGCTGTAAAACCTTTATATGAAACACTATAGTGGTGGTGAAGTCATA 3504
DB 2551 ACTCCATTTTGAAGCTGTAAAACCTTTATATGAAACACTATAGTGGTGGTGAAGTCATA 2610
QY 3505 ATTCTTCACTCCCAAGACTTTTGAAGAAAGGGAAGAAATGATTTGCAGTTATTC 3564
DB 2611 ATTCTTCACTCCCAAGACTTTTGAAGAAAGGGAAGAAATGATTTGCAGTTATTC 2670
QY 3565 GTAATGTCAAGTAGGAGGTATAAAAATATATTGGAATGATCTCTGAAATCCCTGTGGA 3624
DB 2671 GTAATGTCAAAATACCACTTATAAATATATTGGACTGTTATCTCTTGAATCCCTGTGGA 2730
QY 3625 AATCTACATTTGAAGACACATCACTCTGAGTGTATACGCAAAAGAAATTCAGTGAGA 3684
DB 2731 AATCTACATTTGAAGACACATCACTCTGAGTGTATACGCAAAAGAAATTCAGT - AGA 2789
QY 3685 TTATCTTTTAAAGAAACTGTAAAAATAGCAACCACTTATGCACTGATATATTGTAGA 3744
DB 2790 TTATCTTTTAAAGAAACTGTAAAAATAGCAACCACTTATGCACTGATATATTGTAGA 2849
QY 3745 CTGTTTCTCTGTTTATGCTCTTGTGTAATCTACTGACATCATCTTCTCTTGGAT 3804
DB 2850 CTGTTTCTCTGTTTATGCTCTTGTGTAATCTACTGACATCATCTTCTCTTGGAT 2909
QY 3805 AGTGGGTGATAGCAAGTATATTTCTAAAAAACTTTGTAATAAGTTTGTGGCTAAAAAT 3864
DB 2910 AGTGGGTGATAGCAAGTATATTTCTAAAAAACTTTGTAATAAGTTTGTGGCTAAAAAT 2969
QY 3865 GA 3866
||

Db 2970 CA 2971
 RESULT 9
 AAS03025
 ID AAS03025 standard; cDNA; 2980 BP.
 XX
 AC AAS03025;
 XX
 DT 29-AUG-2001 (first entry)
 XX
 DE Human diagnostic and therapeutic (dithp) cDNA sequence #14.
 XX
 KW Human diagnostic and therapeutic molecule; dithp; gene therapy;
 KW thalassemia; cardiovascular disorder; cell proliferative disorder;
 KW cancer; neurodegenerative disorder; autoimmune disorder;
 KW infectious disorder; inflammatory disorder; developmental disorder;
 KW Incyte ID number 3324652dec; intracellular signalling molecule; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200121836-A2.
 XX
 PD 29-MAR-2001.
 XX
 PF 19-SEP-2000; 2000WO-US025643.
 XX
 PR 23-SEP-1999; 99US-0155760P.
 PR 24-SEP-1999; 99US-0155939P.
 PR 24-SEP-1999; 99US-0156294P.
 PR 28-SEP-1999; 99US-0156565P.
 PR 28-SEP-1999; 99US-0156624P.
 PR 28-SEP-1999; 99US-0156625P.
 PR 24-NOV-1999; 99US-0167410P.
 PR 24-NOV-1999; 99US-0167517P.
 PR 24-NOV-1999; 99US-0167520P.
 PR 24-NOV-1999; 99US-0167521P.
 PR 24-NOV-1999; 99US-0167531P.
 PR 24-NOV-1999; 99US-0167542P.
 PR 24-NOV-1999; 99US-0167543P.
 PR 29-NOV-1999; 99US-0167943P.
 PR 29-NOV-1999; 99US-0167945P.
 PR 30-NOV-1999; 99US-0168197P.
 PR 30-NOV-1999; 99US-0168265P.
 PR 30-NOV-1999; 99US-0168429P.
 PR 30-NOV-1999; 99US-0168432P.
 PR 01-DEC-1999; 99US-0168468P.
 PR 01-DEC-1999; 99US-0168599P.
 PR 02-DEC-1999; 99US-0168611P.
 PR 02-DEC-1999; 99US-0168613P.
 PR 02-DEC-1999; 99US-0168857P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Hodgson DM, Lincoln SE, Russo FD, Spiro PA, Banville SC;
 PI Bratcher SR, Dufour GE, Cohen HJ, Rosen BH, Shah P, Chalup MS;
 PI Hillman JL, Jones AL, Yu JV, Greenawalt LB, Panzer SR, Roseberry AM;
 PI Wright RJ, Chen W, Liu TF, Yap PE, Stockdreher TK, Amshey S;
 PI Fong WT;
 XX
 DR WFI; 2001-281607/29.
 XX
 PT Novel diagnostic and therapeutic polynucleotides, used in disease
 PT diagnosis and for gene therapy of conditions such as cancer and
 PT thalassemia.
 XX
 PS Claim 1; Page 260-261; 299pp; English.
 XX
 CC The present sequence for human diagnostic and therapeutic (dithp) cDNA
 CC sequence #14 is 1 of 71 (AAS03012-AAS03082) novel sequences described in
 CC the invention. The present sequence (Incyte ID No: 3324652dec) encodes an
 CC intracellular signalling molecule. The dithp polynucleotides may be used
 CC to diagnose a condition disease or disorder associated with human

CC molecules. They can be used to identify the presence of similar nucleic
 CC acids. Dithp polynucleotides may be used to generate hybridisation probes
 CC for use in chromosomal mapping. Polypeptides (DITHP) encoded by dithp are
 CC used to screen for molecules which bind to them and modulate their
 CC activity. Dithp polynucleotides can be used for gene therapy of disorders
 CC such as severe combined immunodeficiency syndrome (SCID), cystic
 CC fibrosis, thalassemia, haemophilia resulting from Factor VIII or IX
 CC deficiencies, cardiovascular disorders e.g familial hypercholesterolaemia
 CC (FH), cell proliferative disorders e.g. cancers, neurodegenerative
 CC disorders, autoimmune/inflammatory disorders, infectious disorders and
 CC developmental disorders. The antibodies can be used to analyse protein
 CC expression levels
 XX
 SQ Sequence 2980 BP; 1057 A; 527 C; 560 G; 836 T; 0 U; 0 Other;
 Query Match 75.2%; Score 2906; DB 4; Length 2980;
 Best Local Similarity 99.1%; Pred. No. 0;
 Matches 2943; Conservative 0; Mismatches 25; Indels 2; Gaps 2;
 QY 897 GTGAATGTAGACTGTACAGACTTCCTAGAAACAGTTTGGTTCCTATCTTTTCATTTC 956
 DB 2 GCGACGGGAAATCAAAACGTTTGGGAAAGGAGTTGGGTTCCATCTTTTCATTTC 61
 QY 957 CCAGTGCAGTTTCTGTAGAAATCGAATCCGAGGATTTAGTGGCAGAGATTCACAAATT 1016
 DB 62 CCAGGCAG-TTCTGTAGAAATCGAATCCGAGGATTTAGTGGCAGAGATTCACAAATT 120
 QY 1017 GATTCCATAATGAACAAAGTGAGAGACATTAATAAATTAAGTTTAAAAATGAAGACCTTACT 1076
 DB 121 GATTCCATAATGAACAAAGTGAGAGACATTAATAAATTAAGTTTAAAAATGAAGACCTTACT 180
 QY 1077 GATGAACCTAGCTTGAATAAATTTCTGCTGATCTACAGATTAATCTCGGAACCTGTTAAC 1136
 DB 181 GATGAACCTAGCTTGAATAAATTTCTGCTGATCTACAGATTAATCTCGGAACCTGTTAAC 240
 QY 1137 CAATTTATGATGATGGCAACCAACCCAGAGGACTGGTGTAGTTTGTTCCTCAAACTAGAG 1196
 DB 241 CAATTTATGATGATGGCAACCAACCCAGAGGACTGGTGTAGTTTGTTCCTCAAACTAGAG 300
 QY 1197 AAAACAGTGTTCGCTAAGTGTCTCTTTTAAATTAATTAATGTGTGCTTCAAGTCAA 1256
 DB 301 AAAACAGTGTTCGCTAAGTGTCTCTTTTAAATTAATTAATGTGTGCTTCAAGTCAA 360
 QY 1257 GCAATTGAAGCGCTTCCCGCAGATAAATATGGCCAAATCAGAGAGTTTGTCTAGAAATCAA 1316
 DB 361 GCAATTGAAGCGCTTCCCGCAGATAAATATGGCCAAATCAGAGAGTTTGTCTAGAAATCAA 420
 QY 1317 GTGAGATTTGCTGAATTTAAAGCTATTCAAGAGCCAGATGATGACGTGACTACTTTCAA 1376
 DB 421 GTGAGATTTGCTGAATTTAAAGCTATTCAAGAGCCAGATGATGACGTGACTACTTTCAA 480
 QY 1377 ATGGCCAGAGCAAACTCGAAGAAATTTGCTTTGTCATATATCTTTTGTCACAAATTGAA 1436
 DB 481 ATGGCCAGAGCAAACTCGAAGAAATTTGCTTTGTCATATATCTTTTGTCACAAATTGAA 540
 QY 1437 CTGTCAAGAGTAATGTCAAAAAAGTAAACAACTCTTCAAAAAAGCTGTAGAACCTGGA 1496
 DB 541 CTGTCAAGAGTAATGTCAAAAAAGTAAACAACTCTTCAAAAAAGCTGTAGAACCTGGA 600
 QY 1497 GCAGTACACTAGAAATGCTGGAAATTTGCCCTGGGAAATTTAAACCTCCAAAAAAGCAG 1556
 DB 601 GCAGTACACTAGAAATGCTGGAAATTTGCCCTGGGAAATTTAAACCTCCAAAAAAGCAG 660
 QY 1557 CTGCTTTTCAGAGGAGGAGAAAGAGAAATTTATCAGCATCTACCGTATTAACTGCCCAAGAA 1616
 DB 661 CTGCTTTTCAGAGGAGGAGAAAGAGAAATTTATCAGCATCTACCGTATTAACTGCCCAAGAA 720
 QY 1617 TCATTTTCCGGTTCATTGGGCAATTTACAGATAGGAAACAAAGTTGTGATTCAGAGGA 1676
 DB 721 TCATTTTCCGGTTCATTGGGCAATTTACAGATAGGAAACAAAGTTGTGATTCAGAGGA 780
 QY 1677 CAGACTACTAAAGCCAGGTTTATATATGGAGAGAAACATGCCACCAAGATGAGAAATA 1736

RESULT 10	DB	1	ATGGAATCCGAGGATTTAAAGTGGCAGAGAAATGCAATTTGATTCATTAATGAACAAAGTG	60
ABX76453	QY	1038	AGAGACATTAATAAATAAGTTTAAATAATGAAGACCTTACTGATGAACCTTAAGCTTCAATAAAA	1097
ID ABX76453 standard; DNA; 2574 BP.	DB	61	AGAGACATTAATAAATAAGTTTAAATAATGAAGACCTTACTGATGAACCTTAAGCTTCAATAAAA	120
XX AC ABX76453;	QY	1098	ATTTCTGCTGATCTACAGATAACTCGGAACTCTTTAAACCAAAATATGATGATGGCAAAAC	1157
XX DT 02-APR-2003 (first entry)	DB	121	ATTTCTGCTGATCTACAGATAACTCGGAACTCTTTAAACCAAAATATGATGATGGCAAAAC	180
XX DE Lung cancer-associated polynucleotide #317.	QY	1158	AACCCAGAGGACTGGTTCAGTTTCTGCTCAAACTAGAGAAAACAGTGTTCGCTAAGT	1217
XX KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;	DB	181	AACCCAGAGGACTGGTTCAGTTTCTGCTCAAACTAGAGAAAACAGTGTTCGCTAAGT	240
XX KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;	QY	1218	GATGCTCTTTTAAATAAATTTGATTTGCTTTTACAGTCAAGCAATTTGAAGCGCTTCCGCCA	1277
XX KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;	DB	241	GATGCTCTTTTAAATAAATTTGATTTGCTTTTACAGTCAAGCAATTTGAAGCGCTTCCGCCA	300
XX KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;	QY	1278	GATAAATATGSCCAAATGAGAGTTTTTCTGCTAGAAATTCAGTGAGATTTGCTGAATTAATA	1337
XX KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.	DB	301	GATAAATATGSCCAAATGAGAGTTTTTCTGCTAGAAATTCAGTGAGATTTGCTGAATTAATA	360
XX OS Unidentified.	QY	1338	GCTATTCAAGAGCCAGATGATGCACGTGACTTCTTCAAAATGGCCAGAGCAAACTGCAAG	1397
XX FN WO200286443-A2.	DB	361	GCTATTCAAGAGCCAGATGATGCACGTGACTTCTTCAAAATGGCCAGAGCAAACTGCAAG	420
XX PD 31-OCT-2002.	QY	1398	AAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAACTGTCAAGTAATGTCAAA	1457
XX PF 18-APR-2002; 2002WO-US012476.	DB	421	AAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAACTGTCAAGTAATGTCAAA	480
XX PR 18-APR-2001; 2001US-0284770P.	QY	1458	AAAGTAAACAACTCTTCAAAAGCTGTAGACGTGGAGCAGTACCACTAGAAATGCTG	1517
XX PR 10-MAY-2001; 2001US-0290492P.	DB	481	AAAGTAAACAACTCTTCAAAAGCTGTAGACGTGGAGCAGTACCACTAGAAATGCTG	540
XX PR 09-NOV-2001; 2001US-0339245P.	QY	1518	GAAATTTGCCCTGCGGAAATTTAAACCTCCAAAAAAGCAGCTGCTTTTCAGAGGAGGAAAG	1577
XX PR 13-NOV-2001; 2001US-0350666P.	DB	541	GAAATTTGCCCTGCGGAAATTTAAACCTCCAAAAAAGCAGCTGCTTTTCAGAGGAGGAAAG	600
XX PR 29-NOV-2001; 2001US-0334370P.	QY	1578	AAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCAATTTCCGGTTCACTTGGG	1637
XX PR 12-APR-2002; 2002US-0372246P.	DB	601	AAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCAATTTCCGGTTCACTTGGG	660
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.	QY	1638	CATTTACAGAATAGGAACAACAGTTGTGATTCAGAGGACAGACTACTAAAGCCAGGTTT	1697
XX PI Aziz N, Murray R;	DB	661	CATTTACAGAATAGGAACAACAGTTGTGATTCAGAGGACAGACTACTAAAGCCAGGTTT	720
XX DR WPI: 2003-093161/08.	QY	1698	TTATATGAGAGAACATGCCACCAACAGATGAGAAATAGTTTACCGGAATTCATTGAGA	1757
XX PS P-PSDB; ABUS6724.	DB	721	TTATATGAGAGAACATGCCACCAACAGATGAGAAATAGTTTACCGGAATTCATTGAGA	780
XX PT Detecting a lung cancer-associated transcript in a cell from a patient	QY	1758	CAAACTCAAAAATAAAACAGTCTATGCCATTTGGAAGAGTCCCAAGTTAACTTCTAAAT	1817
XX PT for treating lung cancer, by contacting a biological sample from the	DB	781	CAAACTCAAAAATAAAACAGTCTATGCCATTTGGAAGAGTCCCAAGTTAACTTCTAAAT	840
XX PT patient with a polynucleotide that exhibits increased or decreased	QY	1818	AGCCAGATTTGATGTGAAGACAGATGATTCAGTTGATCTGTTTATGAAAAGCAAA	1877
XX PS expression in lung cancer.	DB	841	AGCCAGATTTGATGTGAAGACAGATGATTCAGTTGATCTGTTTATGAAAAGCAAA	900
XX Claim 22; Page 433-434; 453pp; English.	QY	1878	ACCTCTAGATCAGATCCGAGATTTGGTGTGCTGGATCTAAACCAAGTGGAAATGAT	1937
XX The invention relates to a method for detecting a lung cancer-associated	DB	901	ACCTCTAGATCAGATCCGAGATTTGGTGTGCTGGATCTAAACCAAGTGGAAATGAT	960
XX transcript in a cell from a patient, comprising contacting a biological	QY	1938	TCCTGTGAAATTAAGAAATTTAAAGTCTGTTTCAAAATAGTCAATTTCAAGGAACCTCTGGT	1997
XX sample from the patient with a polynucleotide that selectively hybridises	DB	961	TCCTGTGAAATTAAGAAATTTAAAGTCTGTTTCAAAATAGTCAATTTCAAGGAACCTCTGGT	1020
XX to a sequence that is at least 80 % identical to a gene that exhibits	QY	1998	TCAGATGAAAAGAGTTCTGAACTTATTATTACTGATTAATCAATCCCTGAAGATAAAGC	2057
XX increased or decreased expression in lung cancer samples. Lung cancer-	DB	1021	TCAGATGAAAAGAGTTCTGAACTTATTATTACTGATTAATCAATCCCTGAAGATAAAGC	1080
XX associated polynucleotides and polypeptides are used for identifying a	QY	2058	GAATCAAGTCTTCTAGCTAAATTAAGAGAGAACTAAAGAGTATCAAGAACACAGAGTTCCA	2117
XX compound that modulates a lung cancer-associated polypeptide, for	DB			
XX inhibiting proliferation of a lung cancer-associated cell to treat lung	QY			
XX cancer in a patient and for treating a mammal having lung cancer by	DB			
XX administering a modulatory compound identified. The methods are useful	QY			
XX for treating lung cancer, such as small cell lung cancer, non-small cell	DB			
XX lung cancer or other benign or precancerous lesions, e.g. atelectasis,	QY			
XX emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,	DB			
XX hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and	QY			
XX bronchiectasis. The genes, polynucleotides and polypeptides are useful	DB			
XX for diagnostic purposes and as targets for screening for therapeutic	QY			
XX compounds that modulate lung cancer, such as antibodies. Sequences	DB			
XX ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the	QY			
XX invention.	DB			
XX Sequence 2574 BP; 926 A; 463 C; 488 G; 597 T; 0 U; 0 Other;	QY			
XX Query Match 66.5%; Score 2569.2; DB 7; Length 2574;	QY			
XX Best Local Similarity 99.9%; Pred. No. 0;	QY			
XX Matches 2571; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	QY			
XX 978 ATGGAATCCGAGGATTTAAAGTGGCAGAGAAATGCAATTTGATTCATTAATGAACAAAGTG	QY			

CC biological sample from the patient with a polynucleotide that selectively
CC hybridizes to a sequence at least 80% identical to any of one of 80
CC nucleic acid sequences given in the specification. The method is useful
CC in diagnosing ovarian cancer and in identifying and using agents and/or
CC targets that inhibit ovarian cancer. The nucleic acid molecule,
CC polypeptide and the antibody may also be used in detecting ovarian
CC cancers, monitoring and early detection of relapse following treatment,
CC monitoring response to therapy, selecting patients for post-operative
CC chemotherapy or radiation therapy, in selecting mode of therapy,
CC determining tumour prognosis, early detection of pre-cancerous lesions,
CC and as vaccines. This sequence corresponds to one of the nucleic acids
CC used for the detection method of the invention.

XX
SQ Sequence 2574 BP; 926 A; 463 C; 488 G; 697 T; 0 U; 0 Other;

Query Match 66.5%; Score 2569.2; DB 9; Length 2574;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2571; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 978 ATGAATCCGAGATTAAAGTCGACAGAAATGACAAATGATTCCTATATGAACAAAGTG 1037
DB 1 ATGGAATCCGAGATTAAAGTCGACAGAAATGACAAATGATTCCTATATGAACAAAGTG 60

QY 1038 AGAGACATTAATAAATAGTTTAAATGAAGACCTTACTGATGAACCTAAGCTTGAATAAA 1097
DB 61 AGAGACATTAATAAATAGTTTAAATGAAGACCTTACTGATGAACCTAAGCTTGAATAAA 120

QY 1098 ATTTCTGCTGATCTACAGATAACTCGGAACTGTTAAACAAATATGATGCGAAAC 1157
DB 121 ATTTCTGCTGATCTACAGATAACTCGGAACTGTTAAACAAATATGATGCGAAAC 180

QY 1158 AACCCAGAGACGTTGAGTTTGTGCTCAAACTAGAGAAACAGTCTTCGCTAAGT 1217
DB 181 AACCCAGAGACGTTGAGTTTGTGCTCAAACTAGAGAAACAGTCTTCGCTAAGT 240

QY 1218 GATGCTCTTTTAAATAAATGATTTGCTGCTTACAGTCAAGCAATTCGAAGCGTTCCCCCA 1277
DB 241 GATGCTCTTTTAAATAAATGATTTGCTGCTTACAGTCAAGCAATTCGAAGCGTTCCCCCA 300

QY 1278 GATAAATATGCGCAAAATGAGAGTTTGTCTAGAAATTCAGTGAATTTGCTGATTAATA 1337
DB 301 GATAAATATGCGCAAAATGAGAGTTTGTCTAGAAATTCAGTGAATTTGCTGATTAATA 360

QY 1338 GCTATTCAGAGCAGATGATGACGCTGACTCTTTCAAATGCCAGAGCAAACTGCAAG 1397
DB 361 GCTATTCAGAGCAGATGATGACGCTGACTCTTTCAAATGCCAGAGCAAACTGCAAG 420

QY 1398 AAATTTGCTTTGTTGATATATCTTTTGCAAAATTTGAATGCTGCAAGAGTAAATGTCAAA 1457
DB 421 AAATTTGCTTTGTTGATATATCTTTTGCAAAATTTGAATGCTGCAAGAGTAAATGTCAAA 480

QY 1458 AAAAGTAAACAACTCTTTCAAAGAGCTGTAGACGTTGGAGCAGTACCCTAGAAATGCTG 1517
DB 481 AAAAGTAAACAACTCTTTCAAAGAGCTGTAGACGTTGGAGCAGTACCCTAGAAATGCTG 540

QY 1518 GAAATTCCTCGGAAATTTAAACCTTCAAAAAGCAGCTGCTTTTCAGAGGAGGAAAG 1577
DB 541 GAAATTCCTCGGAAATTTAAACCTTCAAAAAGCAGCTGCTTTTCAGAGGAGGAAAG 600

QY 1578 AAGAAATTTATCAGATCTACGTTATTAACCTGCCCAAGATCAATTTCCGGTTCACTTGG 1637
DB 601 AAGAAATTTATCAGATCTACGTTATTAACCTGCCCAAGATCAATTTCCGGTTCACTTGG 660

QY 1638 CAITTTACAGAAATAGGAACAACTGTTGATTTCCAGAGGACAGACTACTAAAGCCAGGTTT 1697
DB 661 CAITTTACAGAAATAGGAACAACTGTTGATTTCCAGAGGACAGACTACTAAAGCCAGGTTT 720

QY 1698 TTATATCGAGAGAACTGCGACCAAGATGCGAATAATAGTTTACCGGAATTCATTTGAGA 1757
DB 721 TTATATCGAGAGAACTGCGACCAAGATGCGAATAATAGTTTACCGGAATTCATTTGAGA 780

QY 1758 CAAACTAAACAACTAAACAGTATGCGCCATTTGGAAGAGTCCAGGTTAACTTCTTAAAT 1817

DB 781 CAAACTAAACAACTAAACAGTATGCCCATTGGAAGAGTCCAGTTAACCTTCTAAAT 840

QY 1818 AGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTATGATGAAAGACAA 1877

DB 841 AGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTATGATGAAAGACAA 900

QY 1878 ACCTCTAGATCAGAAATCCGAGATTGGTTGTGCTGATCTAAACCAAGTGAATGAT 1937

DB 901 ACCTCTAGATCAGAAATCCGAGATTGGTTGTGCTGATCTAAACCAAGTGAATGAT 960

QY 1938 TCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCAATTTCAAGGAACCTCTGGT 1997

DB 961 TCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCAATTTCAAGGAACCTCTGGT 1020

QY 1998 TCAGATGAAAGAGTTCTGAACCTTATTACTGATTCAATAACCTTGAAGATAAAGC 2057

DB 1021 TCAGATGAAAGAGTTCTGAACCTTATTACTGATTCAATAACCTTGAAGATAAAGC 1080

QY 2058 GAATCAAGTCTTCTAGCTAAATTTAGAGAACTAAAGAGTATCAAGAACCCAGAGTTCCA 2117

DB 1081 GAATCAAGTCTTCTAGCTAAATTTAGAGAACTAAAGAGTATCAAGAACCCAGAGTTCCA 1140

QY 2118 GAGAGTAACAGAAACAGTGGCAAGCTTAAGAGAAAGTCAAGAGTGTATTAACCCAGAACTCT 2177

DB 1141 GAGAGTAACAGAAACAGTGGCAATCTAAGAGAAAGTCAAGAGTGTATTAACCCAGAACTCT 1200

QY 2178 GCTGATCTTCAATCACTGGCAGATTCGGAGTTAGCCCGAAAGTTAATACAGAGCAG 2237

DB 1201 GCTGATCTTCAATCACTGGCAGATTCGGAGTTAGCCCGAAAGTTAATACAGAGCAG 1260

QY 2238 AAACATACCCTTTTGAGCAACCTGCTTTTCAAGTTTCAAAACAGTCAACCAACATATCA 2297

DB 1261 AAACATACCCTTTTGAGCAACCTGCTTTTCAAGTTTCAAAACAGTCAACCAACATATCA 1320

QY 2298 ACATCTAAATGTTTGACCCCAAAATCTATTGTAAGACACCAAGCAGCAATACCTGGAT 2357

DB 1321 ACATCTAAATGTTTGACCCCAAAATCTATTGTAAGACACCAAGCAGCAATACCTGGAT 1380

QY 2358 GATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAGAAATGACTTCCACCTGTTGTCAG 2417

DB 1381 GATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAGAAATGACTTCCACCTGTTGTCAG 1440

QY 2418 TTGTCAACACCTTATGGCCAAACCTGCTTTTCCAGCAGCAACAGCATCAAAATCTTGCC 2477

DB 1441 TTGTCAACACCTTATGGCCAAACCTGCTTTTCCAGCAGCAACAGCATCAAAATCTTGCC 1500

QY 2478 ACTCCACTTCAAAATTTTACAGGTTTGTAGCATCTTCTTCCAGCAATGATGATTCGTT 2537

DB 1501 ACTCCACTTCAAAATTTTACAGGTTTGTAGCATCTTCTTCCAGCAATGATGATTCGTT 1560

QY 2538 AAAGGAAGAATTTTATCCATATTAAGCAGATAGGAAGTGGAGTTTCAAGCAAGGTTATTT 2597

DB 1561 AAAGGAAGAATTTTATCCATTTTAAAGCAGATAGGAAGTGGAGTTTCAAGCAAGGTTATTT 1620

QY 2598 CAGTGTAAATGAAAGAAACAGATATATGCTATAAATATGTGAATTTAGAGAGCA 2657

DB 1621 CAGTGTAAATGAAAGAAACAGATATATGCTATAAATATGTGAATTTAGAGAGCA 1680

QY 2658 GATAACCAAACTCTTGTATGATTTACCGGAACGAATAGCTTATTGAAATAAATCAACAA 2717

DB 1681 GATAACCAAACTCTTGTATGATTTACCGGAACGAATAGCTTATTGAAATAAATCAACAA 1740

QY 2718 CACAGTGAATGATCTCCGACTTTATGATTAATAATCAAGCAAGTACATCTACATG 2777

DB 1741 CACAGTGAATGATCTCCGACTTTATGATTAATAATCAAGCAAGTACATCTACATG 1800

QY 2778 GTAATGGAGTGTGGAATATTTGATCTTTAATAGTTGGCTTAAAGAAATAAATCCATGAT 2837

DB 1801 GTAATGGAGTGTGGAATATTTGATCTTTAATAGTTGGCTTAAAGAAATAAATCCATGAT 1860

QY 2838 CCATGGGAACCAAGAGTTACTGGAAAAATATGTTAGAGGAGTTTCACAATTCATCAA 2897

DB 1861 CCATGGGAACCAAGAGTTACTGGAAAAATATGTTAGAGGAGTTTCACAATTCATCAA 1920

QY 2938 CATGGCATTTGTCACAGTGATCTTTAAACACAGCTAACTTTCTGTAGTTGATGAATGCTA 2957
 Db |||||
 QY 1921 CATGGCATTTGTCACAGTGATCTTTAAACACAGCTAACTTTCTGTAGTTGATGAATGCTA 1980
 Db |||||
 QY 2958 AAGCTAATGATTTTGGGATTTGCAACCAATGCAACAGATCAACAAAGTGTGTTAAA 3017
 Db |||||
 QY 1981 AAGCTAATGATTTTGGGATTTGCAACCAATGCAACAGATCAACAAAGTGTGTTAAA 2040
 Db |||||
 QY 3018 GATTCTCAGGTTGGCAGCAGTTAAATATATGTCACCAAGCAATCAAGATATGTTCTTC 3077
 Db |||||
 QY 2041 GATTCTCAGGTTGGCAGCAGTTAAATATATGTCACCAAGCAATCAAGATATGTTCTTC 2100
 Db |||||
 QY 3078 TCCAGAGAAATGGGAATCTAAGTCAAGATAAGCCCCCAAAAGTGATGTTTGGTCCCTTA 3137
 Db |||||
 QY 2101 TCCAGAGAAATGGGAATCTAAGTCAAGATAAGCCCCCAAAAGTGATGTTTGGTCCCTTA 2160
 Db |||||
 QY 3138 GGATGTAATTTGTACTATGACTTACGGGAACACCAATTCAGCAGATAATTAATCAG 3197
 Db |||||
 QY 2161 GGATGTAATTTGTACTATGACTTACGGGAACACCAATTCAGCAGATAATTAATCAG 2220
 Db |||||
 QY 3198 ATTCTAAATPACATGCCATAATGATCCTAATCATGAATTTGAATTTCCCGATATTCCA 3257
 Db |||||
 QY 2221 ATTCTAAATPACATGCCATAATGATCCTAATCATGAATTTGAATTTCCCGATATTCCA 2280
 Db |||||
 QY 3258 GAGAAAGATCTTCAGATGTTGTTAAAGTGTGTTTAAAGGGACCCAAACAGAGATA 3317
 Db |||||
 QY 2281 GAGAAAGATCTTCAGATGTTGTTAAAGTGTGTTTAAAGGGACCCAAACAGAGATA 2340
 Db |||||
 QY 3318 TCCATTCCTGAGTCTCCGGGTCATCCATATGTTCAAAATTCAAACTCATCCAGTTAAACCAA 3377
 Db |||||
 QY 2341 TCCATTCCTGAGTCTCCGGGTCATCCATATGTTCAAAATTCAAACTCATCCAGTTAAACCAA 2400
 Db |||||
 QY 3378 ATGCCAAGGGAACCACTGAAGAAATGAATATGTTCTGGGCAACTGTTGTTGCTGTAAT 3437
 Db |||||
 QY 2401 ATGCCAAGGGAACCACTGAAGAAATGAATATGTTCTGGGCAACTGTTGTTGCTGTAAT 2460
 Db |||||
 QY 3438 TCTCCTAACTCCATTTTGAAGTGTCTAAACCTTTATATGACACACTATAGTGGTGAA 3497
 Db |||||
 QY 2461 TCTCCTAACTCCATTTTGAAGTGTCTAAACCTTTATATGACACACTATAGTGGTGAA 2520
 Db |||||
 QY 3498 AGTCATAATTTCTCATCCTCCAGACTTTTGAAGGGAAGGGGAAAAAATCA 3551
 Db |||||
 QY 2521 AGTCATAATTTCTCATCCTCCAGACTTTTGAAGGGAAGGGGAAAAAATCA 2574
 Db |||||

RESULT 12

ADC37168
 ID ADC37168 standard; DNA; 2860 BP.
 AC
 XX ADC37168;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 1.
 XX
 XX Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
 KW cancer; infectious disease; bone disease; AIDS;
 KW neurodegenerative disease; ischemic disease; Antiinflammatory;
 KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
 KW Neuroprotective; Nootropic; Cardiant; Gene therapy; murine; gene; ds.
 XX
 OS Mus musculus.
 FN W02003048202-A2.
 XX
 XX 12-JUN-2003.
 PD
 XX 03-DEC-2002; 2002WO-JP012644.
 PF
 XX 03-DEC-2001; 2001JP-0036892.
 PR 05-DEC-2001; 2001US-0335829P.
 PR 03-OCT-2002; 2002JP-00291302.

PR 04-OCT-2002; 2002US-0415769P.
 XX (ASAH) ASAH KASEI KK.
 PA
 XX Matsuda A, Muramatsu S;
 PI
 XX WPI; 2003-505282/47.
 DR P-P8DB; ADC37169.
 XX
 PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
 useful for treating inflammation, autoimmune diseases, cancers,
 PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
 PT ischemic disorders.
 XX
 PS Claim 4; SEQ ID NO 1; 938pp; English.
 XX
 CC The present invention relates to novel proteins and their coding
 sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
 CC kappaB). The proteins and their coding sequences are useful for treating
 CC a disease associated with NF-kappaB activation, such as inflammation,
 CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
 CC neurodegenerative diseases, or ischemic disorders.
 XX
 SQ Sequence 2860 BP; 914 A; 601 C; 628 G; 716 T; 0 U; 1 Other;
 Query Match 42.2%; Score 1633; DB 9; Length 2860;
 Best Local Similarity 76.8%; Pred. No. 0;
 Matches 2192; Conservative 0; Mismatches 560; Indels 103; Gaps 12;
 QY 971 TGTAGAAATGGAATCCGAGGATTTAAAGTGGCAGAGAAATGACAAATGATTCCATAATGAA 1030
 Db |||||
 QY 1031 CAAAGTGAAGACATTTAAATAAAGTTTAAATAAAGAACCTTACTGATGAACCTTAAAGCTT 1090
 Db |||||
 QY 1091 GAATAAATTTCTGCTGATCTACAGATAACTCGGAACCTGTTAACCAATATGATGAT 1150
 Db |||||
 QY 166 GTCTAAATCTGTG-----CGATCACACCGAAACTGTTAACCAAAATATGAGGGT 216
 Db |||||
 QY 1151 GCGAAACAAACCCAGAGGACTGGTGGATTTGTTGCTCAAACTAGAGAAAAACAGTGTTC 1210
 Db |||||
 QY 217 TGGGAACACCCAGAGAACTGGTGAATTTCTGCTGAACCTAGAGAAAAACAGCTCACC 276
 Db |||||
 QY 1211 GCTAAGTGAATGCTCTTTTAAATAAATTAATGTTGCTGTACAGTCAAGCAATGAAGCGCT 1270
 Db |||||
 QY 277 TCTAAATGACGATCTTTTAAATAAAGCTGATTGGTGGTATAGTCAAGCGATTGAAGTACT 336
 Db |||||
 QY 1271 TCCCCAGATAAATATGCCAAATGAGAGTTTGTCTAGAATTCAGTGAGATTGCTGA 1330
 Db |||||
 QY 337 TCCTCCAGATAAATACGCCCAATGAGAGCTTTGCTCGAATACAAATGAGACTTGCTGA 396
 Db |||||
 QY 1331 ATTAAGACTATTCAAGAGCCAGATGATGCACTGACTACTTCAAAATGGCCAGAGCAAA 1390
 Db |||||
 QY 397 ACTAAAGCTATTCAAGAGCTGATGATGCCGCTGACTACTTCCAGATGGCCAGGAAAA 456
 Db |||||
 QY 1391 CTGCAAGAAATTCGCTTTGTTGTCATATCTTTGTCACAAATGCACTGTCACAGGTAA 1450
 Db |||||
 QY 457 CTGCAAGAAATTCGCTTTGTTGTCACAGTATCTTTGTCACAGTTTGAATGCTCTCAAGGCAA 516
 Db |||||
 QY 1451 TGTCAAAAAAGTAAACAACTTCTTCAAAAAAGCTGTAGAACTGGAGCAGTACCACCTAGA 1510
 Db |||||
 QY 517 TCTTAAAAAAGTGGAGCAGCTTCTTCAAAAGCTGTAGAGACTGGGCGGTGCGCTGCA 576
 Db |||||
 QY 1511 AATGCTGAAATTCGCTGCGGAATTTAACTCCAAAAAAGCAGCTGCTTTTCAGAGGA 1570
 Db |||||
 QY 577 GATGCTGGAGACGGCCATGCGTAACTTACCTCCAGAAAAAGCAGCTGCTTCCGGAGGA 636
 Db |||||
 QY 1571 GGAAGAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCATTTTCCGGTTC 1630
 Db |||||
 QY 637 GGACAGAGAGAGTGTGTCAGCATGACAGTACTAGTGTCCCAAGAGCGGTCTTCCAGCTC 696
 Db |||||

1631 ACTTGGGCAATTTACAGAAATAGGAACAAACAGTTGTGATTCAGAGGACAGACTACTAAAGC 1690
1697 ACTTGGAAATGTACAGAAATAGGAGCATCAGCTGTGATTCAGAGGACAGGCTGGGCGAGC 756
1691 CAGGTTTTATATGGAGAGAAATGCCACACAGATGCAGAAATAGGTTACCGGAATTC 1750
757 CAGGTTTTATATGGAGAGAAATGCTCCAAAGATGCCGAAGTGAAGCATCAAAACCC 816
1751 ATTGAGCAAACTAACAAACTAAACAGATCATCCCATTTGGAGAGTCCCAATTAACCT 1810
817 CTTCAAGCAGACTCAGCAGCTAAACGGTCAATCCCTTTGGAAGAGTCCCAATCAATCT 876
1811 TCTAAATAGCCAGATTTGATGTGAGACAGATGATTCAGTTGTACCTTGTTTATGAA 1870
877 TCTAAACAGCCAGATTTCTATGTGAGACAGATAGCTCAGCTGTGACACAGTTTAAACAC 936
1871 AAGACAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCTCGATCTAAACCAAGTGG 1930
937 AAGAAAGGGTCAGGACCGA---CCGAGACGGATTTCTGCCGGCTCCAGACCCAGTGG 993
1931 AATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTTCAAAATAGTCAATTTCAAGGAAC 1990
994 CAGTGATTCCTATGAATGAGAGGTTTAAAGCCCATTTCAAACTATCTATTTGAAAGACTC 1053
1991 TCTGGTGTGAGATGAAAGAGTTCTGAACCTTATTATTCTGATTAATCAATCAACCTGAAGAA 2050
1054 TTTGGTGTCAATGAAAGAGTTCTGNAC---TTATGTCTGATTTATAGCCTTGAAGAG 1110
2051 TAAACCGGAATCAAGTCTTCTAGCTAAATTAAGAGAACTTAAAGATATCAAGAACCGA 2110
1111 TAAACAGATTTCAAG---CTTAAACAAATTTGGAAGAACTTAAG-----CCAGA 1155
2111 GGTTCAGAGAGTAACAGAAACAGTGGCAGCTTAAGAGAACTCAGAGTGTATTAAACCA 2170
1156 GATTGCAAGAAAGGCCCATGAGTGGCAGTCTACAGAAAGCCGAGTGTGTGTCCA 1215
2171 GAATCTGTGCTATTTCAAACTCACTGGCAGATTCGCGAGTTAGCCGCAAAAGTTAATAC 2230
1216 GAACCTGTGCTCTTGCACCCCTGGCGACGTTTCCAGATGTCAACCCGAAAGCTGACAA 1275
2231 AGACGAGAAACATACCACCTTTTGAAGCACTGTCTTTTCCAGTTTCAAAACAGTCCACC 2290
1276 AG-----AGTCAACCC 1287
2291 AATATCAACATCTAAATGGTTTGCACCAAAATCTATTGTGAAGACACCAAGCAGCAATAC 2350
1288 AATATCAGTTCTTAATGGCTTGATCCAAAGTCTGCTTGTGAGACACCTAGTAGCAGCTC 1347
2351 CTTGGATGATATCAGTACGCTGTTTGAACCTCAGTGTGTAAGAAATGACTTCCACCTGC 2410
1348 CTTGGATGATATCAGTAAATGTTTGAAGCTCCAGTGTGAAGAAATGACTTCCACCTGC 1407
2411 TTGTCAGTTGTCAACACTTATGGCAACTGTCTGTTTCCAGCAGCAACAGCATCAAT 2470
1408 CTGTCATCATCAACACTTACAGCCAGCTTGGCCGCTCCAGCAGCAACAGCAGCAGGG 1467
2471 ACTTGCCACTCCACTTCAAAATTTACAGGTTTGTAGCATCTTCTTCAGCAAAATGAATGCAT 2530
1468 ACTCAGCACTCCTCTTCAAGCTTGAGATTTTCCAGGTTCTTCAATCAATGAATGCAT 1527
2531 TTCGGTTAAGGAGAGATTTATTCATATTAAAGCAGATAGGAGTGGAGTTTCAAGCAA 2590
1528 TTCAGTTAAGCGAAGATTTATTCATATTAAAGCAGATAGGAGTGGAGTTTCCAGTAA 1587
2591 GGTATTTCAAGTGTAAATGAAAGAAACAGATATATGCTTATAAAATATGTGAACCTTGA 2650
1588 GGTGTTTCAGGTAATTGAATGAGAAAGAAACAGATAAAGCTATCAATATGTGAACCTGAG 1647
2651 AGAAGCAGATAACCAAACTTGTAGTGTACCGGAAAGAAATAGCTTATTCGATTAACCT 2710
1648 AGACGCCGATAGCCAACTATTGAGAGCTACCGCAACGAGATAGCGTTTGTGAACAACT 1707
2711 AACACACACAGTATGAATCATCCGACTTTTATGATTTATGAATTAACGACCCAGTACAT 2770

1708 ACAGCAACACAGTATGAAGATCATCCGCCCTCTATGATTTATGAATCACCCGAGCATACAT 1767
2771 CTACATGGTAAATGAGTGTGAAATATTGATCTTTAATAGTTGGCTTAAAGAGAAAAATC 2830
1768 CTACATGGTAAATGGAATGTGAAACATTGACCTAAATAGTTGGCTTAAAGAGAAAAATC 1827
2831 CATTTGATCCATGGGAAACGAAGAGTTACTTGAAAAATATGTTAGAGCAGTTTACACAAAT 2890
1828 CATCAATCCATGGGAAACGAAGAGTACTTGAAAAACATGTTGGAGGAGTACACATAAT 1887
2891 CCATCAACATGGCAATGTTTCAACAGTATCTTAAACACAGCTAACTTTCTGATAGTTGATGG 2950
1888 CCATCAGCATGTTATGTTCTATAGTATCTGAAGCCTGCTAACTTTGTGATAGTGACCG 1947
2951 AATGCTAAAGCTAAATGATTTTGGGATGTCAAACACCAATGAACCAAGATCAACAAAGTGT 3010
1948 AATGCTAAAGCTAAATGATTTTGGGATGTCAAACACCAATGACCCAGACACAAAGCAT 2007
3011 TGTAAAGATTTCTCAGGTTGGCAGATTAATTTATATCCACCAAGCAATCAAGATAT 3070
2008 TGTAAAGATTTCTCAGGTTGGCAGATTAATTTATATGCCCCCAGAGCAATCAGAGCAT 2067
3071 GTCTTCTCCAGAGAGAAATGGGAAATCTAAGTCAAGATTAAGCCCCAAAAAGTGTATGTTG 3130
2068 GTCTTCTTCAAGAGAAAAATTCGAAATCAGGACCAAGGTAAGTCCAGAGAGTGTATGTCG 2127
3131 GTCTTCTAGGATGTTATTTGTTACTATATGACTTACGGGAAACACCAATTTTCCAGCAGATAAT 3190
2128 GTCTTGGGGTGCATTTTGTACTACATGACTTATGGGAGGACGCCATTTCCAGCAGATCAT 2187
3191 TAATCAGATTTCTAAATATACATGCCATAATGATCTTAATCATCAAAATGAAATTTCCCGA 3250
2188 CAATCAGGTCCTTAACTGACGCCATAATCAACCTGCTCATGAGATTGAAATTTCCCGA 2247
3251 TATTCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAGAGGACCCAAACA 3310
2248 GATTTCGAAAAAGATCTTCGAGACGTTTAAAGTGTGTTTAAAGTGTGTTTAAAGTGTGTTTAAAG 2307
3311 GAGGATATCCATTCCTGAGTCTCTGCTCATCATATGTTTCAAAATTCAAACTCATCCAGT 3370
2308 GAGATATCTATCCCTGAGCTTCTCACATCCGATGTTTCAAAATTCAGCCCCATCCAGG 2367
3371 TAAACAAATGCCAAGGAAACCACTGAAGAAATGAATATGTTCTGCGGCCCACTTGTGG 3430
2368 CAGCAAAATGGCTAGGAGGAGCACTGATGAATGAATGAATATGTTGTTGGTCAACTTGTGG 2427
3431 TCTGAATCTCCTAACTCCATTTTGAAGCTGCTTAAACTTTTATATCAACACTATAGTGG 3490
2428 TCTGAATCTCCTAACTCCATTTTGAAGAACTGTCAAAAACCTTTGATGAACTGTTAATTTG 2487
3491 TGTGAAAGTCAATAAATTTCTTATCCTCCAGACTTTTGAAGAAAAAGGGGAAAAAATG 3550
2488 TGTGAAAGTCAAGATTTCTTCTGTCATCCAGACTTTTGAAGAAAAAGGAAAAAGTGT 2547
3551 ATTTGCAATTTCTGTAATGTGATAGGAGTATAAATATATGAGCTGTTATATCTCT 3610
2548 ATGCACAGCTACGTACAAACCAAGA-----ACACTGATTTGTTTCTCTGCCATCTCT 2601
3611 TGAATCCCTGTGGAAATCTA-CATTTGAAGACCAACATCACTCTGAAGTGT----- 3660
2602 TGAATCTCTGAGGAAATCTACCAGTTGGAACCAACCTCACCTGATTTTATCAGTTAAAA 2661
3661 --ATCAGCAAAAAAATTCAGTGAGATTTCTTTAAAGAAAACTGTAAAAATAGCAACC 3718
2662 AAACAAAAAACAACAACTTCAGTAGATTTATCCCAAAAGGAAGCTGTAAAGTT---AACC 2718
3719 ACTTATGCACTGTATATTTAGTGTGTTTCT-CTGTTTATGCTCTTGTGTAATC 3777
2719 ACTCATAGCACTGTGTATATTAATTTATAGATTTGCTTTTCTTTTATGCTTTCTGTA 2778
3778 TACTTGACATCATTTTACTCTTGGAAATAGTGGGTG 3812

Db 2779 AATCTGCTAATGTTTACGTTTACAGACAGTGAATG 2813

RESULT 13

ID ADA02894/c standard; DNA; 32404 BP.

XX AC ADA02894;

XX DT 06-NOV-2003 (first entry)

XX XX Human BLR1 carcinoma associated gene, SEQ ID NO:1412.

XX KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast; prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening; gene; ds.

XX OS Homo sapiens.

XX FN WO2003057146-A2.

XX PD 17-JUL-2003.

XX PF 26-DEC-2002; 2002WO-US041414.

XX XX 26-DEC-2001; 2001US-00035832.

XX PA (SAGR-) SAGRES DISCOVERY.

XX PI Morris DW;

XX DR WPI; 2003-587068/55.

XX PT New recombinant nucleic acid encoding carcinoma associated protein, useful for preparing compositions for treating carcinomas.

XX PS Claim 1; SEQ ID NO 1412; 245pp; English.

XX CC The invention relates to recombinant carcinoma associated (CA) nucleic acid sequences from mouse and human (ADA01482-ADA03094), and to recombinant carcinoma associated proteins (CAP) encoded by them. The invention also encompasses expression vectors and host cells comprising a CA nucleic acid, a polypeptide (especially an antibody) that specifically binds to the protein, and a biochip comprising CA nucleic acid or fragments thereof. The sequences of the invention were identified using oncogenic retroviruses, which insert into the genome of the host organism at random. Many of these do not carry transduced host oncogenes or pathogenic trans-acting viral genes, meaning that cancer incidence is a direct consequence of the effects of proviral integration into host protooncogenes. The CA nucleic acid sequences can be used to diagnose carcinoma (especially breast cancer, prostate cancer, lymphoma or leukaemia) or a propensity to carcinoma by determination of the sequence of a CA gene, or by determination of CA gene expression in particular tissues. CA nucleic acids, proteins and antibodies are also useful as therapeutic agents and in screening and evaluating drug candidates. The present sequence represents a specifically claimed human CA nucleic acid sequence of the invention. Note: The complete sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

XX SQ Sequence 32404 BP; 7711 A; 8273 C; 9097 G; 7303 T; 0 U; 20 Other;

Query March 10.2%; Score 395.6; DB 8; Length 32404; Best Local Similarity 81.3%; Pred. No. 4e-93; Matches 499; Conservative 0; Mismatches 104; Indels 11; Gaps 3;

Qy 9 TTTTCTTTTGTGAGATGAGTTTCACTCTTTGTTGGCAGGCTGGAGTGAATGCCA 68
Db 31974 TATTATTATTTTGAATCGAGTTTCACTCTTTGTTGGCAGGCTGGAGTGAATGCCA 31915

Qy 69 CAATCTCAGCTTACTGCAACCTCCGCTCCGGGTTCAAGCGATTCTCTGCTCAGCT 128
|||||

Db 31914 CGAACTCGGCTTACTGCAACCTCTGCTCTCTGAGTTCAAGGATTTTCTGCTCAGGCT 31855
Qy 129 CTCAGTAGCTGGGATTACAGGCATGTGCCACCCCTGGCTAACTAATTTCTTTCTTA 188
Db 31854 CCGAGTAGCTGGGATTACAGGCATGTGCCACCCCTGGCTAACTAATTTCTTTCTTA 31800
Qy 189 TTTAGTAGAGATGGGTTTCAACATGTTGGTTCAGGCTGGTCTTTGAATCTCTGCTCAGG 248
Db 31799 TTTAGCAGAGTGGAGTTTTAGCATGTTGGCCAGACCGTCTCTGAATCTCTGACCTCAGG 31740
Qy 249 TGATCCACTTGGCTTGGCTCCCAAGTGTAGGATTACAGCCGTGAACCTGTGCTGGC 308
Db 31739 TGATACACTGCTCAGCCCCCAAGTGTGGGATTACAGCGTGAAGCC--ACCCGGCC 31682
Qy 309 TGATCTCTTTTGTGTTGGATTTTGAACAGAGGT----CTCCCTTTGGTGGCCAGGC 364
Db 31681 TGATCTCTTTTGTGTTGGATTTTGAACAGAGGT----CTCCCTTTGGTGGCCAGGC 31622
Qy 365 TGGAGTGCAGTGGTGGATCTTGGCTCACTATAAATCTCACTCTCTGTTTCAAGTGATC 424
Db 31621 TGGAGTGCAGTGGTGGATCTTGGCTCACTATAAATCTCACTCTCTGTTTCAAGTGATC 31562
Qy 425 CTCCCACTTTAGCTCTGAGTAGTGTGATTACAGGGGTGCCACACACACCCGGCTAA 484
Db 31561 CTCTGCTTACGCTTCCCAAGTGTGGGATTACAGGCATGGTCAACACGCTGGCTAA 31502
Qy 485 TTTTGTATTTTATTAGACAGAGGTTTCAACATGTTGGCCAGGCTGTCTCAAACTGCC 544
Db 31501 TTTTGTATTTTATTAGTAGAGTGGAGTTTCAACATGTTGGCCAGGCTGTCTTGAACCTCC 31442
Qy 545 TGGACTCAAGGATCGGCTGCTCCACTTCCCAAGTCCCGAGATTACAGGTGTGAGTC 604
Db 31441 TGGCTCAAGTATCGCCGCTGGCTCCCAAGTGTGGGATTACAGGCATGAAGCC 31382
Qy 605 ACCATGCTGACCT 618
Db 31381 ACCGTGCCGGCT 31368

RESULT 14

ID ADB72632/c standard; DNA; 32404 BP.

XX AC ADB72632;

XX DT 04-DEC-2003 (first entry)

XX XX Human BLR1 gene.

XX DE human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas; cancer; neoplasm; adenocarcinoma; sarcoma; gene.

XX KW Homo sapiens.

XX OS WO2003008583-A2.

XX PN 30-JAN-2003.

XX PD 26-DEC-2001; 2001WO-US051291.

XX PF 02-MAR-2001; 2001US-00798586.

XX PR 23-OCT-2001; 2001US-00004113.

XX PR 08-NOV-2001; 2001US-00052482.

XX PR 30-NOV-2001; 2001US-00997722.

XX PR 20-DEC-2001; 2001US-00034650.

XX XX (SAGR-) SAGRES DISCOVERY.

XX PA Morris DW, Engelhard EK;

XX PI WPI; 2003-239337/23.

XX DR New recombinant nucleic acid, useful for treating carcinomas, lymphomas,

PT

PT cancers, neoplasm, adenocarcinoma, or sarcomas.
 XX Claim 1; SEQ ID NO 460; 2304pp; English.
 XX
 CC The invention relates to a novel recombinant nucleic acid comprising a
 CC nucleotide sequence selected from any of the 660 sequences fully defined
 CC in the specification. A polynucleotide of the invention has cytostatic
 CC activity, and may have a use in gene therapy, or in a vaccine. The
 CC recombinant nucleic acids and polypeptides are useful for treating
 CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
 CC sarcomas. The present sequence represents a human gene of the invention.
 XX
 SQ Sequence 32404 BP; 7711 A; 8273 C; 9097 G; 7303 T; 0 U; 20 Other;
 Query Match 10.2%; Score 395.6; DB 9; Length 32404;
 Best Local Similarity 81.3%; Pred. No. 4e-83;
 Matches 499; Conservative 0; Mismatches 104; Indels 11; Gaps 3;
 3Y 9 TTTTITTTTTTTTTTGTAGATGGAGTTTCACTCTTTTGGCCAGGCTGGAGTCAATGGCA 68
 Db 31974 TATTATTTATTTTGAATCGAGTTTCACTCTTTTGGCCAGGCTGGAGTCAATGGCA 31915
 Qy 69 CAATCTCAGCTTACTGCAACTCGCTCCCGGTTCAAGCGATTCTCCCTCAGCCT 128
 Db 31914 CGAACTCGGCTTACTGCAACTCGCTCCCGTCTGAGTTCAAGCGATTCTCCCTCAGCCT 31855
 Qy 129 CTCAGTAGCTGGGATTTACAGGATGTGCCACACCCCTGGCTAATTTCTTTCTA 188
 Db 31854 CCCAGTAGCTGGGATTTACAGGATGTGCCACACCCCTGGCTAATTTCTTTCTA 188
 Qy 189 TTTAGTAGAGATGGGGTTTACCAATGTGTGTCAGGCTGGTCTTGAACCTCCTGACCTCAGG 248
 Db 31799 TTTAGCAGAGTGGAGTTTATAGCATGTGGCCAGACCGGTCCTGACCTCAGG 31740
 Qy 249 TGATCCACTTCCCTTGGCTCCCAAGTGTAGATTTACAGCGTGAACCTGCTGCTGGC 308
 Db 31739 TGATACACCTCCCTCAGCGCCCAAGTGTGGGATTTACAGCGTGAACCTGCTGCTGGC 31682
 Qy 309 TGATTTCTTTTGTGTGTGGATTTTGAACACAGGGT----CTCCCTTGTGTCGCCAGGC 364
 Db 31681 TGATTTCTTTTGTGTGTGGATTTTGAACACAGGGT----CTCCCTTGTGTCGCCAGGC 31622
 Qy 365 TGGAGTGCAGTGGCGATCTTGGCTCACTATTAACCTTCCACCTCTGCTGGTTCAAGTGATC 424
 Db 31621 TGGAGTGCAGTGGCGACAATCTCGCTCACTGCAACCTCCACCTCTGCTGGTTCAAGTAAT 31562
 Qy 425 CTCACACTTTAGCTCTCTGAGTAGCTGTGATTACAGCGGTGCACACACACCGGGTAA 484
 Db 31561 CTCCTGCTTCAAGCTCTCCCAAGTAGCTGGGATTTACAGCATGCGTCAACACCGCTGGCTAA 31502
 Qy 485 TTTTGTATTTTATTTAGACACAGGGTTTCAACATGTGTGGCCAGGCTGTCTCAAACTCC 544
 Query Match 10.2%; Score 395.6; DB 9; Length 32404;
 Best Local Similarity 81.3%; Pred. No. 4e-83;
 Matches 499; Conservative 0; Mismatches 104; Indels 11; Gaps 3;
 3Y 9 TTTTITTTTTTTTTTGTAGATGGAGTTTCACTCTTTTGGCCAGGCTGGAGTCAATGGCA 68
 Db 31974 TATTATTTATTTTGAATCGAGTTTCACTCTTTTGGCCAGGCTGGAGTCAATGGCA 31915
 Qy 69 CAATCTCAGCTTACTGCAACTCGCTCCCGGTTCAAGCGATTCTCCCTCAGCCT 128
 Db 31914 CGAACTCGGCTTACTGCAACTCGCTCCCGTCTGAGTTCAAGCGATTCTCCCTCAGCCT 31855
 Qy 129 CTCAGTAGCTGGGATTTACAGGATGTGCCACACCCCTGGCTAATTTCTTTCTA 188
 Db 31854 CCCAGTAGCTGGGATTTACAGGATGTGCCACACCCCTGGCTAATTTCTTTCTA 188
 Qy 189 TTTAGTAGAGATGGGGTTTACCAATGTGTGTCAGGCTGGTCTTGAACCTCCTGACCTCAGG 248
 Db 31799 TTTAGCAGAGTGGAGTTTATAGCATGTGGCCAGACCGGTCCTGACCTCAGG 31740
 Qy 249 TGATCCACTTCCCTTGGCTCCCAAGTGTAGATTTACAGCGTGAACCTGCTGCTGGC 308
 Db 31739 TGATACACCTCCCTCAGCGCCCAAGTGTGGGATTTACAGCGTGAACCTGCTGCTGGC 31682
 Qy 309 TGATTTCTTTTGTGTGTGGATTTTGAACACAGGGT----CTCCCTTGTGTCGCCAGGC 364
 Db 31681 TGATTTCTTTTGTGTGTGGATTTTGAACACAGGGT----CTCCCTTGTGTCGCCAGGC 31622
 Qy 365 TGGAGTGCAGTGGCGATCTTGGCTCACTATTAACCTTCCACCTCTGCTGGTTCAAGTGATC 424
 Db 31621 TGGAGTGCAGTGGCGACAATCTCGCTCACTGCAACCTCCACCTCTGCTGGTTCAAGTAAT 31562
 Qy 425 CTCACACTTTAGCTCTCTGAGTAGCTGTGATTACAGCGGTGCACACACCGGGTAA 484
 Db 31561 CTCCTGCTTCAAGCTCTCCCAAGTAGCTGGGATTTACAGCATGCGTCAACACCGCTGGCTAA 31502
 Qy 545 TGGACTCAAGGATCCCGCTCCCTCACTTCCCAAGTCCAGAGTCCAGAGTATAGAGTGTGAGTC 604
 Db 31441 TGCCCTCAAGTGTATCCCGCCCTCGCCCTTCCCAAGTGTAGGATTTACAGGATTAAGCC 31382
 Qy 605 ACCATGCTCAGCT 618
 Db 31381 ACCGTGCCCGCCT 31368
 RESULT 15
 ID ADC85373/C
 XX standard; DNA; 32404 BP.
 AC
 AC ADC85373;
 XX
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Mouse Blr1 coding sequence.
 XX

Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA; secreted; transmembrane; intracellular; ds.
 Mus sp.
 WO2003045230-A2.
 05-JUN-2003.
 02-DEC-2002; 2002WO-US038592.
 30-NOV-2001; 2001US-00997722.
 (SAGR-) SAGRES DISCOVERY.
 Morris DW, Engelhard BK;
 WPI; 2003-513603/48.
 New recombinant nucleic acid comprising a nucleotide sequence of any of the carcinoma-associated (CA) genes, useful for screening for drug candidates for diagnosing or treating carcinomas.
 Claim 1; SEQ ID NO 159; 983pp; English.
 The invention relates to a recombinant nucleic acid comprising a nucleotide sequence selected from any of the fully defined carcinoma-associated (CA) genes from the 50 tables given in the specification. The CA proteins are secreted, transmembrane or intracellular proteins. The recombinant nucleic acids are useful for screening for drug candidates for diagnosing or treating carcinomas. Sequences given in ADC85215-ADC85514 represent CA genes of the invention.
 Sequence 32404 BP; 7711 A; 8273 C; 9097 G; 7303 T; 0 U; 20 Other;
 Query Match 10.2%; Score 395.6; DB 9; Length 32404;
 Best Local Similarity 81.3%; Pred. No. 4e-83;
 Matches 499; Conservative 0; Mismatches 104; Indels 11; Gaps 3;
 3Y 9 TTTTITTTTTTTTTTGTAGATGGAGTTTCACTCTTTTGGCCAGGCTGGAGTCAATGGCA 68
 Db 31974 TATTATTTATTTTGAATCGAGTTTCACTCTTTTGGCCAGGCTGGAGTCAATGGCA 31915
 Qy 69 CAATCTCAGCTTACTGCAACTCGCTCCCGGTTCAAGCGATTCTCCCTCAGCCT 128
 Db 31914 CGAACTCGGCTTACTGCAACTCGCTCCCGTCTGAGTTCAAGCGATTCTCCCTCAGCCT 31855
 Qy 129 CTCAGTAGCTGGGATTTACAGGATGTGCCACACCCCTGGCTAATTTCTTTCTA 188
 Db 31854 CCCAGTAGCTGGGATTTACAGGATGTGCCACACCCCTGGCTAATTTCTTTCTA 188
 Qy 189 TTTAGTAGAGATGGGGTTTACCAATGTGTGTCAGGCTGGTCTTGAACCTCCTGACCTCAGG 248
 Db 31799 TTTAGCAGAGTGGAGTTTATAGCATGTGGCCAGACCGGTCCTGACCTCAGG 31740
 Qy 249 TGATCCACTTCCCTTGGCTCCCAAGTGTAGATTTACAGCGTGAACCTGCTGCTGGC 308
 Db 31739 TGATACACCTCCCTCAGCGCCCAAGTGTGGGATTTACAGCGTGAACCTGCTGCTGGC 31682
 Qy 309 TGATTTCTTTTGTGTGTGGATTTTGAACACAGGGT----CTCCCTTGTGTCGCCAGGC 364
 Db 31681 TGATTTCTTTTGTGTGTGGATTTTGAACACAGGGT----CTCCCTTGTGTCGCCAGGC 31622
 Qy 365 TGGAGTGCAGTGGCGATCTTGGCTCACTATTAACCTTCCACCTCTGCTGGTTCAAGTGATC 424
 Db 31621 TGGAGTGCAGTGGCGACAATCTCGCTCACTGCAACCTCCACCTCTGCTGGTTCAAGTAAT 31562
 Qy 425 CTCACACTTTAGCTCTCTGAGTAGCTGTGATTACAGCGGTGCACACACCGGGTAA 484
 Db 31561 CTCCTGCTTCAAGCTCTCCCAAGTAGCTGGGATTTACAGCATGCGTCAACACCGCTGGCTAA 31502
 Qy 485 TTTTGTATTTTATTTAGACACAGGGTTTCAACATGTGTGGCCAGGCTGTCTCAAACTCC 544

Db	31501	TTTTTGTAATTTTAGTAGAGATGGAGTTTCACCAATTATGCCAGGCTGTCTTGAACCC	31442
Qy	545	TGGACTCAAGGATCCGCTGCCTCCACTTCCCAAGTCCCGAGATTACAGGTGTGAGTC	604
Db	31441	TGCCCTCAAGTGAATCCGCCCGCTCGGCTCCCAAGTGTAGGATTACAGGCATAAGCC	31382
Qy	605	ACCATGCCTGACCT	618
Db	31381	ACCGTGCCCGGCT	31368

Search completed: September 30, 2004, 02:36:07
 Job time : 1334 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 16:02:29 ; Search time 134 Seconds
(without alignments)
1773.302 Million cell updates/sec

Title: US-10-081-119-14
Perfect score: 4376
Sequence: 1 MNKVRDINKKFKNEDLTDEL.....GGESHNSSSSKTFKKRGKK 841

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4376	100.0	841	5	ABP54938 Human tyr
2	4376	100.0	841	7	ABU61611 Human Tyr
3	4376	100.0	841	7	ADD49934 Human lun
4	4376	100.0	841	7	AGE40437 Human TTK
5	4376	100.0	841	7	AGE31645 Human 168
6	4376	100.0	857	7	AGE38369 Human pro
7	4373	99.9	857	6	ABU56724 Lung canc
8	4373	99.9	857	7	ABU80470 Ovarian c
9	4373	99.9	857	7	ADC37171 Nuclear f
10	3182.5	72.7	830	7	ADC37169 Nuclear f
11	673.5	15.4	678	5	ABP54948 Schizosac
12	673.5	15.4	678	7	ABU61621 Yeast MPH
13	647.5	14.8	777	3	AG38500 Arabidops
14	643	14.7	720	3	AG38501 Arabidops
15	640.5	14.6	638	3	AG38502 Arabidops
16	637	14.6	630	4	ABU63414 Drosophil
17	603	13.8	735	6	ABJ25382 Aspergill
18	603	13.8	842	6	ABJ25382 Aspergill
19	586.5	13.4	764	5	ABP54947 Saccharom
20	586.5	13.4	764	7	ABU61620 Yeast MPS
21	585.5	13.4	764	2	AAW26727 Yeast Mps
22	558	12.8	690	5	ABP73673 Candida a
23	523.5	12.0	331	3	AG05720 Arabidops
24	490	11.2	297	3	AG05721 Arabidops
25	472	10.8	282	3	AG05722 Arabidops

26 342.5 7.8 749 7 ADC37183
27 342.5 7.8 809 4 AAM41064
28 340 7.8 713 6 ABU07446
29 338 7.7 528 5 ABP273734
30 338 7.7 717 7 ADC37185
31 338 7.7 718 4 AAM39278
32 336 7.7 1090 3 AAB10946
33 332.5 7.6 250 5 ABB78798
34 332 7.6 598 7 ADC37181
35 332 7.6 809 6 ABU11737
36 323.5 7.5 749 5 ABU65050
37 323.5 7.5 787 7 ADE38387
38 323.5 7.5 748 4 ABG26847
39 325.5 7.4 787 2 AAR76061
40 324 7.4 520 4 ABB64044
41 323.5 7.4 588 4 AAB65659
42 319 7.3 231 4 AAG67541
43 319 7.3 540 2 AAR76062
44 318.5 7.3 431 3 AAB27933
45 318.5 7.3 541 3 AAB27995

ALIGNMENTS

RESULT 1
ABP54938
ID ABP54938 standard; protein; 841 AA.
XX
AC ABP54938;
XX
XX 13-JAN-2003 (first entry)
XX
XX Human tyrosine threonine kinase.
XX
KW Tyrosine threonine kinase; TTK; protein kinase; enzyme; human;
KW Mitotic checkpoint; colon cancer; breast cancer; cytostatic;
KW gene therapy.
XX
OS Homo sapiens.
XX
XX WO200268444-A1.
XX
XX 06-SEP-2002.
XX
XX 21-FEB-2002; 2002WO-US005278.
XX
XX 21-FEB-2001; 2001US-0271254F;
XX
XX (CHIR) CHIRON CORP.
XX
XX Reinhard C, Jefferson AB, Chan VW;
XX
XX WPI; 2002-698650/75.
XX
XX N-PSDB; ABV73974.
XX
XX Reducing growth of cancer cells comprises reducing Tyrosine Threonine
XX Kinase (TTK) activity, useful in diagnosing and treating disorders with
XX abnormal expression levels and activity of TTK, such as lung, colon,
XX prostate and ovarian cancer.
XX
XX Claim 7; Page 71-73; 113pp; English.

XX The present sequence is the protein sequence of human tyrosine threonine
XX kinase (TTK), a mixed specificity protein kinase that is involved in the
XX cell cycle spindle assembly checkpoint. The invention is based on the
XX finding that TTK is differentially expressed in colon tumour cells
XX relative to healthy tumour colon cells as detected by microarray
XX analysis. Differential expression was confirmed in cell lines derived
XX from various forms of cancer, indicating involvement of TTK in cancer as a
XX more general mechanism. Disruption of TTK function using antisense
XX oligonucleotides decreased proliferation, inhibited anchorage-independent
XX growth, and induced apoptosis of cancer cell lines, including a

CC metastatic breast cancer cell line (MDA-MB-213) and a colorectal
 CC carcinoma cell line (SW620). These findings serve as the basis for the
 CC materials and methods of the invention for use in diagnosing cancer of a
 CC patient, for planning and selection of appropriate therapeutic and/or
 CC prophylactic treatment, and for treatment of cancer associated with
 CC aberrant TTK levels, e.g. associated with overexpression or
 CC overproduction of TTK, by inhibition of gene product production (by
 CC decreasing levels of transcription and/or translation), by decreasing TTK
 CC activity by decreasing TTK gene product production, and/or by reducing
 CC one or more of TTK's kinase activities in cancer cells, especially breast
 CC cancer or colon cancer cells. The present invention has the advantage of
 CC having the ability to project disease progression based on expression of
 CC TTK in a malignant or pre-malignant growth. It allows a more systematic
 CC approach for intervention of a cancerous disease
 CC
 CC Sequence 841 AA;
 SQ

Query Match 100.0%; Score 4376; DB 5; Length 841;
 Best Local Similarity 100.0%; Pred. No. 2.1e-292;
 Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKVRDIKNKFKNEDLTDELNLKISADTTDNGTGNQIMNMANNPEDWLSLLKLEKNS 60
 DB |||||
 QY 1 MNKVRDIKNKFKNEDLTDELNLKISADTTDNGTGNQIMNMANNPEDWLSLLKLEKNS 60
 DB |||||
 QY 61 VPLSDALLNLKIGRYSQAIEALPPDKYQONESFARIOVRFAELKAIQEPDDARDYFQMAR 120
 DB |||||
 QY 61 VPLSDALLNLKIGRYSQAIEALPPDKYQONESFARIOVRFAELKAIQEPDDARDYFQMAR 120
 DB |||||
 QY 121 ANCKKFAFVHISFAQFELSQGVKSKQLLOKQAVRGAVPLEMLEIALNLNLQKQLLS 180
 DB |||||
 QY 121 ANCKKFAFVHISFAQFELSQGVKSKQLLOKQAVRGAVPLEMLEIALNLNLQKQLLS 180
 DB |||||
 QY 181 EEEKKNLSASTVLTAAQSFSGSLGHLQNRNNSCDRGOTTKARFLYGENNPPDABIGYR 240
 DB |||||
 QY 181 EEEKKNLSASTVLTAAQSFSGSLGHLQNRNNSCDRGOTTKARFLYGENNPPDABIGYR 240
 DB |||||
 QY 241 NSLRQTNKTQSCPFGRVFNLLNSPDCDVKTDDSVVPCFMKQTSRSCRDLVWPGSKP 300
 DB |||||
 QY 241 NSLRQTNKTQSCPFGRVFNLLNSPDCDVKTDDSVVPCFMKQTSRSCRDLVWPGSKP 300
 DB |||||
 QY 301 SGNDSCBLNLKSVQNSHFKPEVSDKSSSELIITDITLKNKTESLLAKLEENKEYOE 360
 DB |||||
 QY 301 SGNDSCBLNLKSVQNSHFKPEVSDKSSSELIITDITLKNKTESLLAKLEENKEYOE 360
 DB |||||
 QY 361 PEVPESNQKQWAKRKSECINQNPAASSNHQIPELARKVNTQEQHTTPEQPVFSVSKQS 420
 DB |||||
 QY 361 PEVPESNQKQWAKRKSECINQNPAASSNHQIPELARKVNTQEQHTTPEQPVFSVSKQS 420
 DB |||||
 QY 421 PP1STSKWDPKSICTPSSNTLDDYMSCFRTPVVKNDFPPACQLSTPYGQACFQQQOH 480
 DB |||||
 QY 421 PP1STSKWDPKSICTPSSNTLDDYMSCFRTPVVKNDFPPACQLSTPYGQACFQQQOH 480
 DB |||||
 QY 481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFVNLNEKKOIIAYKVN 540
 DB |||||
 QY 481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFVNLNEKKOIIAYKVN 540
 DB |||||
 QY 541 LEADNQTLDYSYNEAYLNKLOQHSKDIIRLYDYBITQYIYMWECNIDNLSWLKXK 600
 DB |||||
 QY 541 LEADNQTLDYSYNEAYLNKLOQHSKDIIRLYDYBITQYIYMWECNIDNLSWLKXK 600
 DB |||||
 QY 601 KSIDPWRKSYWKNMLEAVHTTHQIGIVHSDLKPNANFLIVDGMKLIDFGIANQMOPDTT 660
 DB |||||
 QY 601 KSIDPWRKSYWKNMLEAVHTTHQIGIVHSDLKPNANFLIVDGMKLIDFGIANQMOPDTT 660
 DB |||||
 QY 661 SVYKDSQVGTVMYMPPEALKDMSSSRENGSKSKISPKSDVMWSLGCILYYMTYGTPTFQ 720
 DB |||||
 QY 661 SVYKDSQVGTVMYMPPEALKDMSSSRENGSKSKISPKSDVMWSLGCILYYMTYGTPTFQ 720
 DB |||||
 QY 721 IINQISKLHAIDPNHEIEFPDIPKDLQVLKCLKRPDKORISIPELLAHPYVQIOIOTH 780
 DB |||||
 QY 721 IINQISKLHAIDPNHEIEFPDIPKDLQVLKCLKRPDKORISIPELLAHPYVQIOIOTH 780
 DB |||||

QY 781 PYNQMAKGTTERMKYVLGQLVGLNSPNSILKAAKTLTYEHYSGGSHNSSSKTTEKRGK 840
 DB |||||
 QY 781 PYNQMAKGTTERMKYVLGQLVGLNSPNSILKAAKTLTYEHYSGGSHNSSSKTTEKRGK 840
 DB |||||
 QY 841 K 841
 DB 841 K 841

RESULT 2

ABU61611

ID ABU61611 standard; protein; 841 AA.

XX AC ABU61611;

XX DT 11-AUG-2003 (first entry)

XX DE Human Tyrosine threonine kinase, TTK.

XX KW Human; enzyme: tyrosine threonine kinase; TTK; cancer; cytostatic;

XX KM mitotic checkpoint gene.

XX OS Homo sapiens.

XX PN US2003045491-A1.

XX PD 06-MAR-2003.

XX PF 21-FEB-2002; 2002US-00081119.

XX PR 23-FEB-2001; 2001US-0289813P.

XX PA (REIN/) REINHARD C.

XX PA (JEFF/) JEFFERSON A. B.

XX PA (CHAN/) CHAN V W.

XX PI Reinhard C, Jefferson AB, Chan VW;

XX DR WPI; 2003-456566/43.

XX DR N-PSDB; ACA62263.

XX PT Detecting cancer in a subject, by comparing expression levels of tyrosine

XX PT threonine kinase polypeptide or polynucleotide in a subject cell and a

XX PT normal cell, where an increase in the expression level in the test cell

XX PT is indicative of cancer.

XX PS Claim 7; Page 25-27; 79pp; English.

XX CC The invention relates to detecting cancer (other than ovarian cancer) in

XX CC a subject, comprising comparing the expression levels of tyrosine

XX CC threonine kinase (TTK, a mitotic checkpoint gene) polypeptide or

XX CC polynucleotide in a test cell obtained from the subject and in a normal

XX CC non-cancer cell, where an increase in the expression level of TTK protein

XX CC or nucleic acid in the test cell compared to that in the normal cell,

XX CC indicates the presence of cancer other than ovarian cancer. Also included

XX CC are reducing growth of a cancerous cell (by contacting a cancerous cell

XX CC with an amount of an agent effective to reduce TTK polypeptide activity

XX CC in the cell), an assay for identifying a candidate agent that reduces

XX CC growth of a cancerous cell (comprising: (i) detecting the activity of a

XX CC TTK polypeptide in the presence of a candidate agent; and (ii) comparing

XX CC the activity of TTK polypeptide in the presence of a candidate agent

XX CC relative to TTK polypeptide activity in the presence of a candidate

XX CC agent), identifying an agent that reduces TTK activity (comprising: (i)

XX CC contacting a cancerous cell displaying elevated expression of a TTK-

XX CC encoding polynucleotide with a candidate agent; and (ii) determining the

XX CC effect of the candidate agent on TTK polypeptide activity) and assessing

XX CC the prognosis of a cancerous disease other than ovarian cancer in a

XX CC subject (comprising: (i) detecting expression of TTK -encoding

XX CC polynucleotide in a test cancer cell of a subject; and (ii) comparing a

XX CC level of expression of TTK-encoding polynucleotide in the test cancer

XX CC cell with a level of expression of the polynucleotide in a control non-

XX CC cancer cell, where the level of expression of TTK in the test cancer cell

XX CC relative to the level of expression in the control non-cancer cell is

XX CC

CC indicative of the prognosis of the cancerous disease). The methods are
CC useful for detecting cancer (other than ovarian cancer) in a subject,
CC reducing growth of cancerous cells, identifying a candidate agent that
CC reduces growth of a cancerous cell, identifying an agent that reduces TTK
CC activity and assessing the prognosis of a cancerous disease other than
CC ovarian cancer. The methods are also useful for determining the ability
CC of a subject to respond to a particular therapy e.g. as a basis of
CC rational therapy. The present sequence represents human TTK
XX
SQ Sequence 841 AA;

Query Match 100.0%; Score 4376; DB 7; Length 841;
Best Local Similarity 100.0%; Pred. No. 2,1e-292; Indels 0; Gaps 0;
Matches 841; Conservative 0; Mismatches 0;

QY 1 MNKVRDIKKNFKNEDLTDELSLNKI SADTTDSGTVNQIMMANNPEDWLSLLKLEKNS 60
Db 1 MNKVRDIKKNFKNEDLTDELSLNKI SADTTDSGTVNQIMMANNPEDWLSLLKLEKNS 60

QY 61 VPLSDALLNKLI GRYSAIEALPPDKYQNESFARIQVFAELKAI QEPDDARDYFQMAR 120
Db 61 VPLSDALLNKLI GRYSAIEALPPDKYQNESFARIQVFAELKAI QEPDDARDYFQMAR 120

QY 121 ANCKKFAFVHSIQAPELSQGNVKKSKQLLOKXAVRGAVPLEMLEIALRNLNKQKQLLS 180
Db 121 ANCKKFAFVHSIQAPELSQGNVKKSKQLLOKXAVRGAVPLEMLEIALRNLNKQKQLLS 180

QY 181 EEEKXNLASATVLTQAESFSGSLGHLQNRNNSCDRSGQTTKARFLYGENMPPDABIGYR 240
Db 181 EEEKXNLASATVLTQAESFSGSLGHLQNRNNSCDRSGQTTKARFLYGENMPPDABIGYR 240

QY 241 NSLRQTNKTKQSCFPGFVFNLLNSPDCVKTDDSVVPCFMKQTSRSECRDLVVPQSKP 300
Db 241 NSLRQTNKTKQSCFPGFVFNLLNSPDCVKTDDSVVPCFMKQTSRSECRDLVVPQSKP 300

QY 301 SGNDSCELRNLKSVQNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360
Db 301 SGNDSCELRNLKSVQNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360

QY 361 PEVPESNOKQWAKRKSECINQNPAAASNHWQIPELARKNTQKHTTPEQVPSVSKQS 420
Db 361 PEVPESNOKQWAKRKSECINQNPAAASNHWQIPELARKNTQKHTTPEQVPSVSKQS 420

QY 421 PPISTSKWFDPKSICKTPESSNTLDYMSCFRTPVVKNDFPPACOLSTPYQCPACFQOOQH 480
Db 421 PPISTSKWFDPKSICKTPESSNTLDYMSCFRTPVVKNDFPPACOLSTPYQCPACFQOOQH 480

QY 481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSKVPQVNLNEKKQIYAIKYN 540
Db 481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSKVPQVNLNEKKQIYAIKYN 540

QY 541 LEEADNQTLDYSRNEIAYLNKLOQHSDKILRLDYDEITDOYIYMWECGNDILNSWLKXK 600
Db 541 LEEADNQTLDYSRNEIAYLNKLOQHSDKILRLDYDEITDOYIYMWECGNDILNSWLKXK 600

QY 601 KSIDPWERKSYWKNMLEAVTHIHQIGVHSDLPANFLIVDGMKLIIDFGIANQMOPDTT 660
Db 601 KSIDPWERKSYWKNMLEAVTHIHQIGVHSDLPANFLIVDGMKLIIDFGIANQMOPDTT 660

QY 661 SVWKDSQVGTNYNPPPEAIKDMSSRENGKSKISPKSDVWSIGCLLYNYTYGKTBPQQ 720
Db 661 SVWKDSQVGTNYNPPPEAIKDMSSRENGKSKISPKSDVWSIGCLLYNYTYGKTBPQQ 720

QY 721 IINQISKLHAIIDPNHETEFDPIDKLDQVLKCKLRDPKQISIPPELLAHPVVIQTH 780
Db 721 IINQISKLHAIIDPNHETEFDPIDKLDQVLKCKLRDPKQISIPPELLAHPVVIQTH 780

QY 781 PVNQMAKGTTEEMKVVQLGOLVGLNSPNSILKAAKTLIYHYGGGSHNSSSSKTEKRGK 840
Db 781 PVNQMAKGTTEEMKVVQLGOLVGLNSPNSILKAAKTLIYHYGGGSHNSSSSKTEKRGK 840

QY 841 K 841

Db 841 K 841

RESULT 3
ADD49934
ID ADD49934 standard; protein; 841 AA.
XX
AC ADD49934;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human lung specific tumour antigen L 1479p.
XX
KW Human; lung cancer antigen; cytostatic; lung cancer; gene therapy;
XX
OS Homo sapiens.
XX
PN US2003194764-A1.
XX
PD 16-OCT-2003.
XX
PF 04-APR-2002; 2002US-00116712.
XX
PR 05-APR-2001; 2001US-0282289P.
XX
PR 05-OCT-2001; 2001US-0327511P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Bangur CS, Switzer A;
XX
DR WPI; 2003-844452/78.
XX
DR N-PSDB; ADD49931.
XX
PT New isolated polypeptides and polynucleotides useful for diagnosing,
XX preventing and treating cancer, particularly lung cancer.
XX
PS Example 3; SEQ ID NO 666; 250pp; English.
XX
CC The invention relates to an isolated polynucleotide (a) comprising any of
CC the 666 fully defined nucleotide sequences appearing as ADD49269 -
CC ADD49935, ADD49936 and ADD49939, complements of (a); sequences of at
CC least 20 contiguous residues of (a); sequences that hybridise to (a)
CC under highly stringent conditions; sequences having at least 75 or 90%
CC identity to (a); or degenerate variants of (a). Also included are an
CC isolated polypeptide (b) (comprising: sequences encoded by the new
CC polynucleotide; any of the 4 amino acid sequences fully defined in the
CC specification; or sequences having at least 70 or 90% identity to the
CC sequence in (a) or (b)), an expression vector comprising the above
CC polynucleotide operably linked to an expression control sequence, a host
CC cell transformed or transfected with the above expression vector, an
CC isolated antibody, or its antigen-binding fragment, that specifically
CC binds to the above polypeptide, an oligonucleotide that hybridises to the
CC above-mentioned nucleotide sequences under highly stringent conditions, a
CC fusion protein comprising at least one polypeptide cited above, detecting
CC the presence of a cancer in a patient (comprising: obtaining a biological
CC sample from the patient; contacting the biological sample with a binding
CC agent that binds to the polypeptide, or with the oligonucleotide cited
CC above; detecting in the sample an amount of the polypeptide that binds to
CC the binding agent, or an amount of a polynucleotide that hybridises to
CC the oligonucleotide; and comparing the amount of polypeptide, or
CC polynucleotide that hybridises to the oligonucleotide, to a predetermined
CC cut-off value and then determining the presence of a cancer in the
CC patient), a method for stimulating and/or expanding T-cells specific for
CC a tumour protein (comprising contacting T-cells with the above
CC polypeptide, polynucleotide or antigen-presenting cells that express the
CC polynucleotide, under conditions and for a time sufficient to permit the
CC stimulation and/or expansion of T-cells), an isolated T-cell population
CC comprising T-cells prepared by the method, a composition comprising a
CC first component selected from physiological carriers and
CC immunostimulants, and a second component selected from the above
CC polypeptide, polynucleotide, antibody, fusion protein, T-cell population
CC and antigen-presenting cells that express the above polypeptide,
CC

Query Match	100.0%;	Score 4376;	DB 7;	Length 841;
Best Local Similarity	100.0%;	Pred. No. 2.1e-292;		
Matches 841;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MNKVRDIKNKFKNEDELTDLSLNKISADTTDNGSTVQNMANNPDWLSLLKLEKNS	60	
Db	1	MNKVRDIKNKFKNEDELTDLSLNKISADTTDNGSTVQNMANNPDWLSLLKLEKNS	60	
QY	61	VPLSDALLNKILGRYSQAIEALPDKYQONESFARIOVFAELKAQEPDDADYFQMAR	120	
Db	61	VPLSDALLNKILGRYSQAIEALPDKYQONESFARIOVFAELKAQEPDDADYFQMAR	120	
QY	121	ANCKKFAFVHISFAQFELSQGNVKKSKOLLOKQAVERGAVPLEMLEIALRNINLQKKQLLS	180	
Db	121	ANCKKFAFVHISFAQFELSQGNVKKSKOLLOKQAVERGAVPLEMLEIALRNINLQKKQLLS	180	
QY	181	BEEXKNLSASTVLTQASFSGLHQLNRNNSCDRSGQTTKARFLYGENMPPQDAIGYR	240	
Db	181	BEEXKNLSASTVLTQASFSGLHQLNRNNSCDRSGQTTKARFLYGENMPPQDAIGYR	240	
QY	241	NSLRQTNKTKQSCFFGRVVPVNLNNSPCDVKTDDSVVPCFMKQTSRSECRDLVVPQSKP	300	
Db	241	NSLRQTNKTKQSCFFGRVVPVNLNNSPCDVKTDDSVVPCFMKQTSRSECRDLVVPQSKP	300	
QY	301	SGNDSCELRNLKSYONSHFKEPLVSDKSELIITDITLKNKTESLLAKLEETKEYOE	360	
Db	301	SGNDSCELRNLKSYONSHFKEPLVSDKSELIITDITLKNKTESLLAKLEETKEYOE	360	
QY	361	PEVPESKQKQAKKSECINQNPAASSHQWQPELARKVNTQKHTTTEQPVFSVSKOS	420	
Db	361	PEVPESKQKQAKKSECINQNPAASSHQWQPELARKVNTQKHTTTEQPVFSVSKOS	420	
QY	421	PPISTSKWFDPKSICKTPSSNTLDDYVSCFRTPVKNDFPPACQLSTPYGPACFOQOQH	480	
Db	421	PPISTSKWFDPKSICKTPSSNTLDDYVSCFRTPVKNDFPPACQLSTPYGPACFOQOQH	480	
QY	481	QILATPLONQVLASSANECISVKGRIYSILKQIGSGSSKVFQVNLNKKQIYAIKYVN	540	
Db	481	QILATPLONQVLASSANECISVKGRIYSILKQIGSGSSKVFQVNLNKKQIYAIKYVN	540	
QY	541	LEADNQTLDYRNEIAYLNKLOHSDKIIRLDYDEITDQIYVMMECGNIDLNSWLKKK	600	
Db	541	LEADNQTLDYRNEIAYLNKLOHSDKIIRLDYDEITDQIYVMMECGNIDLNSWLKKK	600	
QY	601	KSIDPWERKSYWKNMLEAVHTIHCHGIVHSDLPANFLIVDGMKLIDFGIANQMPDPT	660	
Db	601	KSIDPWERKSYWKNMLEAVHTIHCHGIVHSDLPANFLIVDGMKLIDFGIANQMPDPT	660	
QY	661	SVVKDSQVGTVMYMPPEAIKDMSSRENGSKSIKPSKDVMSLGCILYMTYKTPFQO	720	
Db	661	SVVKDSQVGTVMYMPPEAIKDMSSRENGSKSIKPSKDVMSLGCILYMTYKTPFQO	720	
QY	721	IINQISKLHAIIDPNHEIEPDIPEKDQVLKCCLPKQRIPIPELLAHPPYVQIQTH	780	
Db	721	IINQISKLHAIIDPNHEIEPDIPEKDQVLKCCLPKQRIPIPELLAHPPYVQIQTH	780	
QY	781	PVNMAGTTEEMKYVLGQVLGNSPNSILKAATLYEHYSGESHNSSSSKTFEKKRK	840	
Db	781	PVNMAGTTEEMKYVLGQVLGNSPNSILKAATLYEHYSGESHNSSSSKTFEKKRK	840	
QY	841	K 841		
Db	841	K 841		

RESULT 6

ADE38369

ID ADE38369 standard; protein; 857 AA.

XX AC ADE38369;

XX AC ADE38369;

DT 29-JAN-2004 (first entry)

XX	Human protein 1682 amino acid sequence.
XX	tumorigenic disorder; angiogenic disorder; aberrant gene expression;
XX	aberrant protein activity; cytostatic; antithyroid; antidiabetic;
KW	ophthalmological; cancer; breast cancer; colon cancer; lung cancer;
KW	prostatic cancer; Grave's disease; diabetic retinopathy; protein 1682.
XX	Homo sapiens.
XX	WO2003065006-A2.
XX	07-AUG-2003.
PD	30-JAN-2003; 2003WO-US002588.
XX	31-JAN-2002; 2002US-0353600P.
PR	15-MAR-2002; 2002US-0364517P.
PR	09-APR-2002; 2002US-0371075P.
PR	10-APR-2002; 2002US-0371507P.
PR	16-APR-2002; 2002US-0372984P.
PR	19-APR-2002; 2002US-0374194P.
PR	24-MAY-2002; 2002US-0382995P.
PR	31-MAY-2002; 2002US-0385023P.
PR	14-JUN-2002; 2002US-0388853P.
PR	17-JUN-2002; 2002US-0389395P.
PR	25-JUN-2002; 2002US-0391324P.
PR	15-JUL-2002; 2002US-0395944P.
PR	22-JUL-2002; 2002US-0397726P.
PR	13-AUG-2002; 2002US-0403046P.
PR	22-AUG-2002; 2002US-0405155P.
PR	27-AUG-2002; 2002US-0406361P.
PR	25-OCT-2002; 2002US-0421195P.
PR	12-NOV-2002; 2002US-0425456P.
PR	19-NOV-2002; 2002US-0427626P.
PR	10-DEC-2002; 2002US-0432122P.
XX	(MILL-) MILLENNIUM PHARM INC.
PA	Hunter JJ, Macbeth KJ, Tsai F, Lesoon A, Lightcap ES;
XX	Williamson MW, Rudolph-Owen LA;
PI	WPI; 2003-646176/61.
XX	N-FSDB; ADE38368.
DR	Treating subject having tumorigenic disorder or angiogenic disorder
DR	caused by aberrant polypeptide e.g., N-formylpeptide receptor or nucleic
DR	acid, by administering a modulator.
XX	Disclosure; SEQ ID NO 30; 454pp; English.
PS	This invention relates to a novel method of treating a human subject
XX	having a tumorigenic disorder or angiogenic disorder, caused by aberrant
CC	gene expression or activity of an isolated protein, by administering a
CC	modulator. The modulator may have cytostatic, antithyroid, antidiabetic
CC	or ophthalmological activity. The method is useful for treating a subject
CC	having a tumorigenic or angiogenic disorder, in particular for treating
CC	cancer (for example breast cancer, colon cancer, lung cancer or prostatic
CC	cancer) and, for example, Grave's disease and diabetic retinopathy. The
CC	present sequence is the amino acid sequence of the novel isolated human
CC	protein 1682 of the invention.
XX	Sequence 857 AA;
SQ	Query Match 100.0%; Score 4376; DB 7; Length 857;
	Best Local Similarity 100.0%; Pred. No. 2.1e-292;
	Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MNKVRDIKNKFKNEDELTDLSLNKISADTTDNGSTVQNMANNPDWLSLLKLEKNS 60
Db	17 MNKVRDIKNKFKNEDELTDLSLNKISADTTDNGSTVQNMANNPDWLSLLKLEKNS 76
QY	61 VPLSDALLNKILGRYSQAIEALPDKYQONESFARIOVFAELKAQEPDDADYFQMAR 120

Db	77	VFLSDALLNKLIGYSQAIEALPPDKYGNESFARIQVRFALKAIQEPDADYFQMAR	136
Qy	121	ANCKKFAFVHISFAQFELSGNWKSKKQLQKAVERGAVPLEMIEIALRNLNQKQLLS	180
Db	137	ANCKKFAFVHISFAQFELSGNWKSKKQLQKAVERGAVPLEMIEIALRNLNQKQLLS	196
Qy	181	EEKKNLASVTLTAQSFSGSLGHLQNRNNSCDRSGQTTKARFLYGENMPPQDAEIGYR	240
Db	197	EEKKNLASVTLTAQSFSGSLGHLQNRNNSCDRSGQTTKARFLYGENMPPQDAEIGYR	256
Qy	241	NSLRQTNKTKQSCPFGRVPVNLNNSPCDVKTDSDVVPFCMKRQTSRSECRDLVVPQSKP	300
Db	257	NSLRQTNKTKQSCPFGRVPVNLNNSPCDVKTDSDVVPFCMKRQTSRSECRDLVVPQSKP	316
Qy	301	SGNDSCELRNLKSVQNSHFKEPLYSDEKSSLEIITDSITLKNKTESLLAKLEETKEYQE	360
Db	317	SGNDSCELRNLKSVQNSHFKEPLYSDEKSSLEIITDSITLKNKTESLLAKLEETKEYQE	376
Qy	361	PEVESKQKQAKREKSECINONPAASNHWQIPELARKNTEQKHTTFEQPVFSVSKQS	420
Db	377	PEVESKQKQAKREKSECINONPAASNHWQIPELARKNTEQKHTTFEQPVFSVSKQS	436
Qy	421	PPISTSKWDFPKSICKTPSSNTLDDYMSCRTPVVKNDFFPACQLSTPYGQPAQFQOOQH	480
Db	437	PPISTSKWDFPKSICKTPSSNTLDDYMSCRTPVVKNDFFPACQLSTPYGQPAQFQOOQH	496
Qy	481	QILATPLQNLQVLASSANECISVKGRISYILKQIGSGSKVPQVLEKQYAIKYN	540
Db	497	QILATPLQNLQVLASSANECISVKGRISYILKQIGSGSKVPQVLEKQYAIKYN	556
Qy	541	LEADNOTLDSYRNEIAYLNKLOQHSKIRLYDYEITDQIYVWMECGNIDLSNMLKXX	600
Db	557	LEADNOTLDSYRNEIAYLNKLOQHSKIRLYDYEITDQIYVWMECGNIDLSNMLKXX	616
Qy	601	KSIDPWERKSWKMLSAVHTIHQGIIVHSDLPANFLIVDGMKLIDFGIANQMOPDTT	660
Db	617	KSIDPWERKSWKMLSAVHTIHQGIIVHSDLPANFLIVDGMKLIDFGIANQMOPDTT	676
Qy	661	SVVKDSQVGTVMPPPAIKDMSSSRENGSKSKISPKSDVSIWGLCITYMTYKTPFQ	720
Db	677	SVVKDSQVGTVMPPPAIKDMSSSRENGSKSKISPKSDVSIWGLCITYMTYKTPFQ	736
Qy	721	IINQISKLHAIIDENHIEFPDPEKDLQDVLKCLXRDPKQIRISPELLAHYVQIQTH	780
Db	737	IINQISKLHAIIDENHIEFPDPEKDLQDVLKCLXRDPKQIRISPELLAHYVQIQTH	796
Qy	781	PVQWAKGTTTEMKYVLGOLVGLNSPNSILKAAKTLHYHSGGSHNSSSSKTFEKKRGK	840
Db	797	PVQWAKGTTTEMKYVLGOLVGLNSPNSILKAAKTLHYHSGGSHNSSSSKTFEKKRGK	856
Qy	841	K 841	
Db	857	K 857	
RESULT 7			
ABU56724	ID	ABU56724 standard; protein; 857 AA.	
XX	XX		
AC	ABU56724;		
XX	XX		
DT	02-APR-2003	(first entry)	
DE	Lung cancer-associated polypeptide #317.		
KW	Lung cancer-associated polypeptide; cytostatic; emphysema;		
KW	antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;		
KW	small cell lung cancer; benign lesion; precancerous lesion; bronchitis;		
KW	chronic obstructive pulmonary disease; hypersensitivity pneumonitis;		
XX	interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.		
OS	Unidentified.		

XX	WO200286443-A2.		
PN	31-OCT-2002.		
PD	18-APR-2002; 2002WO-US012476.		
XX	18-APR-2001; 2001US-0284770P.		
PR	10-MAY-2001; 2001US-0290492P.		
PR	09-NOV-2001; 2001US-0339245P.		
PR	13-NOV-2001; 2001US-0350666P.		
PR	29-NOV-2001; 2001US-0334370P.		
PR	12-APR-2002; 2002US-0372246P.		
XX	(EOSB-) EOS BIOTECHNOLOGY INC.		
PA	Aziz N, Murray R;		
XX	WPI: 2003-093161/08.		
DR	N-PSDB; ABX76453.		
DR	Detecting a lung cancer-associated transcript in a cell from a patient		
PT	for treating lung cancer, by contacting a biological sample from the		
PT	patient with a polynucleotide that exhibits increased or decreased		
PT	expression in lung cancer.		
XX	Claim 27; Page 434; 453pp; English.		
PS	The invention relates to a method for detecting a lung cancer-associated		
CC	transcript in a cell from a patient, comprising contacting a biological		
CC	sample from the patient with a polynucleotide that selectively hybridises		
CC	to a sequence that is at least 80 % identical to a gene that exhibits		
CC	increased or decreased expression in lung cancer samples. Lung cancer-		
CC	associated polynucleotides and polypeptides are used for identifying a		
CC	compound that modulates a lung cancer-associated polypeptide, for		
CC	inhibiting proliferation of a lung cancer-associated cell to treat lung		
CC	cancer in a patient and for treating a mammal having lung cancer by		
CC	administering a modulatory compound identified. The methods are useful		
CC	for treating lung cancer, such as small cell lung cancer, non-small cell		
CC	lung cancer or other benign or precancerous lesions, e.g. atelectasis,		
CC	emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,		
CC	hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and		
CC	bronchiectasis. The genes, polynucleotides and polypeptides are useful		
CC	for diagnostic purposes and as targets for screening for therapeutic		
CC	compounds that modulate lung cancer, such as antibodies. Sequences		
CC	ABU56408-ABU56745 represent lung cancer-associated polypeptides of the		
CC	invention		
XX	Sequence 857 AA;		
SQ	Query Match 99.9%; Score 4373; DB 6; Length 857;		
	Best Local Similarity 99.9%; Pred. No. 3.4e-292;		
	Matches 840; Conservative 1; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	MKKVRDINKKFKNEIDLTDLSLNKISADTTDSNGTVNQIMMANNPDWLSLLLEKNS	60
Db	17	MKKVRDINKKFKNEIDLTDLSLNKISADTTDSNGTVNQIMMANNPDWLSLLLEKNS	76
Qy	61	VFLSDALLNKLIGYSQAIEALPPDKYGNESFARIQVRFALKAIQEPDADYFQMAR	120
Db	77	VFLSDALLNKLIGYSQAIEALPPDKYGNESFARIQVRFALKAIQEPDADYFQMAR	136
Qy	121	ANCKKFAFVHISFAQFELSGNWKSKKQLQKAVERGAVPLEMIEIALRNLNQKQLLS	180
Db	137	ANCKKFAFVHISFAQFELSGNWKSKKQLQKAVERGAVPLEMIEIALRNLNQKQLLS	196
Qy	181	EEKKNLASVTLTAQSFSGSLGHLQNRNNSCDRSGQTTKARFLYGENMPPQDAEIGYR	240
Db	197	EEKKNLASVTLTAQSFSGSLGHLQNRNNSCDRSGQTTKARFLYGENMPPQDAEIGYR	256
Qy	241	NSLRQTNKTKQSCPFGRVPVNLNNSPCDVKTDSDVVPFCMKRQTSRSECRDLVVPQSKP	300
Db	257	NSLRQTNKTKQSCPFGRVPVNLNNSPCDVKTDSDVVPFCMKRQTSRSECRDLVVPQSKP	316

Db 617 KSIDPWERKSWKNNLEAVHTIHQHGIVHSDLPANFLIVDGLMKLIDFGIANQMPDPT 676

Qy 661 SVVKDSQVGTVMYMPPEAIKDMSSRENGSKSKISPKSDVWSLGCILYMTYGTPTFQ 720

Db 677 SVVKDSQVGTVMYMPPEAIKDMSSRENGSKSKISPKSDVWSLGCILYMTYGTPTFQ 736

Qy 721 IINQISKHLAIDPNHEIEFPDIPEKDQVVKCCLKRPKQISIPELLAHPYVQIOTH 780

Db 737 IINQISKHLAIDPNHEIEFPDIPEKDQVVKCCLKRPKQISIPELLAHPYVQIOTH 796

Qy 781 PVNQAKGTTEEMKYVLQVLGNSPNSILKAAKTLYEHYSGSHSSSSKTFEKKRGK 840

Db 797 PVNQAKGTTEEMKYVLQVLGNSPNSILKAAKTLYEHYSGSHSSSSKTFEKKRGK 856

Qy 841 K 841

Db 857 K 857

RESULT 9

ID ADC37171 standard; protein; 857 AA.

XX AC ADC37171;

XX DT 18-DEC-2003 (first entry)

XX DE Nuclear factor kappa B (NF-kappaB) activating protein, SEQ ID 4.

XX KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease; cancer; infectious disease; bone disease; AIDS;

XX KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;

XX KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-Hiv;

XX KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human.

XX OS Homo sapiens.

XX WN WO2003048202-A2.

XX PD 12-JUN-2003.

XX PF 03-DEC-2002; 2002WO-JP012644.

XX PR 03-DEC-2001; 2001JP-00368692.

XX PR 05-DEC-2001; 2001US-0335829P.

XX PR 03-OCT-2002; 2002JP-00291302.

XX PR 04-OCT-2002; 2002US-0415769P.

XX PA (ASAH) ASAH KASEI KK.

XX PI Matsuda A, Muramatsu S;

XX DR WPI; 2003-505282/47.

XX DR N-PSDB; ADC37170.

XX PT New purified protein that activates nuclear factor kappa B (NF-kappaB), useful for treating inflammation, autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS, neurodegenerative diseases or ischemic disorders.

XX PS Claim 1; SEQ ID NO 4; 938pp; English.

XX CC The present invention relates to novel proteins and their coding sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-kappaB). The proteins and their coding sequences are useful for treating a disease associated with NF-kappaB activation, such as inflammation, autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS, neurodegenerative diseases, or ischaemic disorders.

XX SQ Sequence 857 AA;

Query Match 99.9%; Score 4373; DB 7; Length 857;

Best Local Similarity 99.9%; Pred. No. 3.4e-292;

Matches 840; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNKVRDIKNKFNEDLTDELINLKISADTTNSGTVQIMMMANNPDMWLSLLLEKNS 60

Db 17 MNKVRDIKNKFNEDLTDELINLKISADTTNSGTVQIMMMANNPDMWLSLLLEKNS 76

Qy 61 VPLSDALLINKLIGRYSQAIEALPPDKYQGNESFARIQVRFAELKAIQEPDDARDYFQAR 120

Db 77 VPLSDALLINKLIGRYSQAIEALPPDKYQGNESFARIQVRFAELKAIQEPDDARDYFQAR 136

Qy 121 ANCKKFAFVHISFAQFELSQGNVKSQKLLQKAVERGAVPLEMLEIALNLNKKQLLS 180

Db 137 ANCKKFAFVHISFAQFELSQGNVKSQKLLQKAVERGAVPLEMLEIALNLNKKQLLS 196

Qy 181 EEEKKNLSASTVLTAEPSFSGSLHLQNRNNSCDRGQTTKARFLYGNMPPQDAEIGYR 240

Db 197 EEEKKNLSASTVLTAEPSFSGSLHLQNRNNSCDRGQTTKARFLYGNMPPQDAEIGYR 256

Qy 241 NSLROTKTKOSCPGRVPVNLNSPCDVKTDSDVWPCFMKROTSECELDLVVPGSKP 300

Db 257 NSLROTKTKOSCPGRVPVNLNSPCDVKTDSDVWPCFMKROTSECELDLVVPGSKP 316

Qy 301 SGNDSCELRNLSQVNSHFKEPLVSDKSELIITDSITLKNKTESSLLAKLEETKEYQE 360

Db 317 SGNDSCELRNLSQVNSHFKEPLVSDKSELIITDSITLKNKTESSLLAKLEETKEYQE 376

Qy 361 PEVPESNOKOWAKKESKINCPNPAASNNHWOIPELARKVNTQKHHTFEQVFSVSKOS 420

Db 377 PEVPESNOKOWAKKESKINCPNPAASNNHWOIPELARKVNTQKHHTFEQVFSVSKOS 436

Qy 421 PPSTSKWFDPKSIKCTPSSNTLDDYMSCFRTPVVKNDPPACQLSTPYGQACFQQOQH 480

Db 437 PPSTSKWFDPKSIKCTPSSNTLDDYMSCFRTPVVKNDPPACQLSTPYGQACFQQOQH 496

Qy 481 QILATPQLQVLAASSANECSIVKGRISYILKQIGSGSSKVPQVLAINEKQIYAIKYN 540

Db 497 QILATPQLQVLAASSANECSIVKGRISYILKQIGSGSSKVPQVLAINEKQIYAIKYN 556

Qy 541 LEEADNQTLDSYRNEIAYLNKLQHSDKIIRLYDYETDQIYVMWECGNIDLSNWLKK 600

Db 557 LEEADNQTLDSYRNEIAYLNKLQHSDKIIRLYDYETDQIYVMWECGNIDLSNWLKK 616

Qy 601 KSIDPWERKSWKNNLEAVHTIHQHGIVHSDLPANFLIVDGLMKLIDFGIANQMPDPT 660

Db 617 KSIDPWERKSWKNNLEAVHTIHQHGIVHSDLPANFLIVDGLMKLIDFGIANQMPDPT 676

Qy 661 SVVKDSQVGTVMYMPPEAIKDMSSRENGSKSKISPKSDVWSLGCILYMTYGTPTFQ 720

Db 677 SVVKDSQVGTVMYMPPEAIKDMSSRENGSKSKISPKSDVWSLGCILYMTYGTPTFQ 736

Qy 721 IINQISKHLAIDPNHEIEFPDIPEKDQVVKCCLKRPKQISIPELLAHPYVQIOTH 780

Db 737 IINQISKHLAIDPNHEIEFPDIPEKDQVVKCCLKRPKQISIPELLAHPYVQIOTH 796

Qy 781 PVNQAKGTTEEMKYVLQVLGNSPNSILKAAKTLYEHYSGSHSSSSKTFEKKRGK 840

Db 797 PVNQAKGTTEEMKYVLQVLGNSPNSILKAAKTLYEHYSGSHSSSSKTFEKKRGK 856

Qy 841 K 841

Db 857 K 857

RESULT 10

ID ADC37169

ID ADC37169 standard; protein; 830 AA.

XX AC ADC37169;

XX DT 18-DEC-2003 (first entry)

XX DE Nuclear factor kappa B (NF-kappaB) activating protein, SEQ ID 2.

XX

KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
 KW cancer; infectious disease; bone disease; AIDS;
 KW neurodegenerative disease; ischaemic disorder; Antinflammatory;
 KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
 KW Neuroprotective; Nootropic; Cardiant; Gene therapy; murine.
 XX Mus musculus.
 XX WO2003048202-A2.
 XX 12-JUN-2003.
 XX 03-DEC-2002; 2002WO-JP012644.
 XX 03-DEC-2001; 2001JP-00368692.
 XX 05-DEC-2001; 2001US-0335829P.
 XX 03-OCT-2002; 2002JP-00291302.
 XX 04-OCT-2002; 2002US-0415769P.
 XX (ASAH) ASahi KASEI KK.
 XX Matsuda A, Muramatsu S;
 XX WPI; 2003-505282/47.
 XX N-PSDB; ADC37169.
 XX New purified protein that activates nuclear factor kappa B (NF-kappaB),
 XX useful for treating inflammation, autoimmune diseases, cancers,
 XX infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
 XX ischemic disorders.
 XX Claim 1; SEQ ID NO 2; 938pp; English.
 XX The present invention relates to novel proteins and their coding
 XX sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
 XX kappaB). The proteins and their coding sequences are useful for treating
 XX a disease associated with NF-kappaB activation, such as inflammation,
 XX autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
 XX neurodegenerative diseases, or ischaemic disorders.
 XX Sequence 830 AA;
 Query Match 72.7%; Score 3182.5; DB 7; Length 830;
 Best Local Similarity 74.0%; Pred. No. 3.3e-210;
 Matches 622; Conservative 87; Mismatches 105; Indels 27; Gaps 7;
 QY 1 MNKVRDINKFKNEDLFDLSLNKISADTTDNGSTVQNMANNPEDWLSLLKLEKNS 60
 DB 17 MSXMRDIKNI-NEDCTDELSLKI CADHTE--TVNQIMRVGNTPENWLNFLKLEKNS 72
 QY 61 VPLSDALLNKLIGRYSOAIALPDKYGONESFARIOVFAELKATQEPDARDYFQMAR 120
 DB 73 SPLNDDLNLNKLIGRYSOAIEVLPPDKYGONESFARIOVRLAELKATQEPDARDYFQMAR 132
 QY 121 ANCKKFAFVHISFAQFELSGNKKQLQLQKAVERGAVPLEMLELALNKLQKKQLLS 180
 DB 133 ENCKKFAFVHVSFAQFELSGNKKSEQLLHKAVETGAVPLOWLETAMENLHLQKKQLLP 192
 QY 181 EEEKKNLSASTVLTQAQSPGSGS;GHLQNRNNSCDSRGQTTKARFLYGENMPQDABIGYR 240
 DB 193 EEDKKSVSASTVLSAQEPFSSGLGNVQNRNISCESRGQAGARVLYGENLPPQDAEVRHQ 252
 QY 241 NSLRQNTKTKQSPFGFVGVNLLNSPDCVDKTDVVPFCMKRQTSRSECDLVVPGSKP 300
 DB 253 NPQKQTHAARKSPFGFVGVNLLNSPDPFYKTDSSAVTQLTTRKSGSPD-RAAILPGSRP 311
 QY 301 SGNDSCBLNKLKQVNSHFKEPLVSKSELSLITDITLXNKTSSSLAKUEPTEYKEQE 360
 DB 312 RGSDSYELRGKPTQITLYLKDLSLVNKEKSEL-MSDLIALKSKTDS-LTKLEETK- --- 365
 QY 361 PEYPESNQKQWQARKKSECNQNPAASSNHQWLPALARKVNTQKHTTEPQPVFSVKOS 420
 DB 366 PEIAERRPMQWQTRKECVFQNPAPAPLRHVVDVTPKAD-----KES 409

QY 421 PPISTKWFDPKSICTKTPSSNTLDYNSCFRTPVVXKNDFFPAQQLSTPYGQACFOQOQH 480
 DB 410 PPIVUPKWLDPKSACTPSSSSLDYWKCKPTFVVKNDFFPACPSSTPYQLARLQOQOQ 469
 QY 481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVPQVNEKKQIYAIKYN 540
 DB 470 QGLSTPLQSLQISGSSSINECISVNGRIYSILKQIGSGSSKVPQVNEKKQINAIKYN 529
 QY 541 LEEADNQLDSYRNEIAYLANKLQOHSDKIRLYDYEITDOYIYVMWECGNDLSNLKXK 600
 DB 530 LEDADSQTIESYRNEIAPLANKLQOHSDKIRLYDYEITEQIYIYVMWECGNDLSNLKXK 589
 QY 601 KSIDPWERKSYWKNMLEAVHTIHQGIHSDLPANFLIVDGMCLKLIDFGIANQMPDIT 660
 DB 590 KSNPWERKSYWKNMLEAVHIIHQGIHSDLPANFVIDGMCLKLIDFGIANQMPDIT 649
 QY 661 SVVKDSQVGTVNYMPPPAIKDMSSSRENGSKSKISPKSDVNSLGCILTYMTYCKTPFQ 720
 DB 650 SIVKDSQVGTVNYMPPPAIRDMSSSRENSKIRTKVSPRSDVNSLGCILTYMTYCKTPFQ 709
 QY 721 IINOISKLHAIIDPNHEIEFPDPEKLODLVKCLKRPKQISIPPELLAHFYVQIQTH 780
 DB 710 IINQVSKLHAIINPAHEIEFPSEKDLRDVLKCLVRNPKERISIPPELLTHEYVQIQPH 769
 QY 781 PVNQMAKTTTEEMKYVLGQVLGNLSPNSILKAAKTLYEHYSGGESHNSSSSKTFEKKRGK 840
 DB 770 PGSQWARGATDEMYKVLGQVLGNLSPNSILKTAKTLYRYNCGEGQDSSSSSKTFDKRRER 829
 QY 841 K 841
 DB 830 K 830
 RESULT 11
 ABP54948
 ID ABP54948 standard; protein; 678 AA.
 XX AC ABP54948;
 XX 13-JAN-2003 (first entry)
 XX Schizosaccharomyces pombe protein kinase Mph1.
 XX Mph1; tyrosine threonine kinase; TTK; protein kinase; enzyme;
 XX mitotic checkpoint; colon cancer; breast cancer; tumour; cytostatic;
 XX gene therapy.
 XX Schizosaccharomyces pombe.
 XX WO200268444-A1.
 XX 06-SEP-2002.
 XX 21-FEB-2002; 2002WO-US005278.
 XX 21-FEB-2001; 2001US-0271254P.
 XX (CHIR) CHIRON CORP.
 XX Reinhard C, Jefferson AB, Chan VW;
 XX WPI; 2002-698650/75.
 XX N-PSDB; ABV73994.
 XX Reducing growth of cancer cells comprises reducing Tyrosine Threonine
 XX Kinase (TTK) activity, useful in diagnosing and treating disorders with
 XX abnormal expression levels and activity of TTK, such as lung, colon,
 XX prostate and ovarian cancer.
 XX Disclosure; Page 104-105; 113pp; English.
 XX The present sequence is the protein sequence of Schizosaccharomyces pombe

[illegible]

Qy 582 IYMWECGNDLNSWLKK-----KSIDPWERKSYKQMLBVAVTHIQGIVHSDLK 633
 Db 494 IYVLEYGEIDLAMLSQKREIEGSDRTIDENLWLFYMQOILQAVNTHEERIVHSDLK 553
 Qy 634 PANFLIVDGLKLIIDFGIAQMOPDTSVVKDSQVGTNVNMPPEAIKDMSSSRENGKSKS 693
 Db 554 PANFLVGRGKLIIDFGIAKAINSDTNIORDSQVGTLSVMSPEAFM-CNESDENGT-I.611
 Qy 694 KISPSVWSLGCILYVNTYKTPFQOIQISKLHAIIDPNHEIFPDIPEKLODVLYK 753
 Db 612 KCGRPSDWSLGCILYQWYGRTPFADYKTFWAKFKVITDPNHEITYNQLSNPWLIDLMK 671
 Qy 754 CCLKRDPKQKISIPPELLAHPV 775
 Db 672 KCLANDRNQWRIPPELLQHPFL 693

RESULT 14

AAG38501

ID AAG38501 standard; protein; 720 AA.

AC AAG38501;

XX XX

DT 18-OCT-2000 (first entry)

XX XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 47507.

XX XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; Genetic mapping; gene expression control; promoter;

KW termination sequence.

XX XX

OS Arabidopsis thaliana.

XX XX

PN EP1033405-A2.

XX XX

PD 06-SEP-2000.

XX XX

FF 25-FEB-2000; 2000EP-00301439.

XX XX

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 28-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132487P.

PR 11-MAY-1999; 99US-0132863P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136382P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139750P.

PR 18-JUN-1999; 99US-0139763P.

PR 21-JUN-1999; 99US-0139817P.

PR 22-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.

PR 24-JUN-1999; 99US-0140695P.

PR 28-JUN-1999; 99US-0140823P.

PR 29-JUN-1999; 99US-0140991P.

PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.

PR 01-JUL-1999; 99US-0142154P.

PR 02-JUL-1999; 99US-0142055P.

PR 06-JUL-1999; 99US-0142390P.

PR 08-JUL-1999; 99US-0142803P.

PR 09-JUL-1999; 99US-0142920P.

PR 12-JUL-1999; 99US-0142977P.

PR 13-JUL-1999; 99US-0143542P.

PR 14-JUL-1999; 99US-0143624P.

PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.

PR 16-JUL-1999; 99US-0144086P.

PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.

PR 19-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.

PR 20-JUL-1999; 99US-0144684P.

PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.

PR 21-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145085P.

PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.

PR 22-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145188P.

PR 23-JUL-1999; 99US-0145224P.

PR 26-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.

PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.

PR 02-AUG-1999; 99US-0146388P.

PR 02-AUG-1999; 99US-0146389P.

PR 03-AUG-1999; 99US-0147038P.

OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX XX 25-FEB-1999; 99US-011895P.
PR 05-MAR-1999; 99US-01231180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127482P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135133P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-01394750P.
PR 18-JUN-1999; 99US-01394763P.
PR 21-JUN-1999; 99US-0139617P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140895P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147132P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149910P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.

PR	15-SEP-1999;	99US-0154018P.	
PR	16-SEP-1999;	99US-0154039P.	
PR	20-SEP-1999;	99US-0154779P.	
PR	22-SEP-1999;	99US-0155133P.	
PR	23-SEP-1999;	99US-0155486P.	
PR	24-SEP-1999;	99US-0155659P.	
PR	28-SEP-1999;	99US-0156458P.	
PR	29-SEP-1999;	99US-0156596P.	
PR	04-OCT-1999;	99US-0157117P.	
PR	05-OCT-1999;	99US-0157753P.	
PR	06-OCT-1999;	99US-0157865P.	
PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158232P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161992P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match			14.6%; Score 640.5; DB 3; Length 638;
Best Local Similarity			29.2%; Pred. No. 2.8e-35;
Matches 187; Conservative 88; Mismatches 218; Indels 147; Gaps 18;			
QY	188	SASTVLTAESEFSGSLGHLNRNNSCDRSGOTTAKRFLYG---ENMPPQDAEIGYRSLR	244
DB	10	SLESLDVQRKSQSLIGTSQDMEDATNQAEASHLDACIGSKHQLPSVDSEVSLKSEYK	69
QY	245	QTNK-TKQSCPFGRVPVNLNNSPCDVKTDDSVVPCFMKQTSRSECDLWFGSKPSGN	303
DB	70	DSSSLAKIQGQIGRFP-NFLNQP-----RTRC-----	95
QY	304	DSCELRNLKSVQSHFKEPLVSDSEKSELITDITLKNKTESLLAKLEETKEYQPEV	363
DB	96	-----SAVGSWATTTLIHSSAPMLNAT-----THVRSYVEAD-130	
QY	364	PESNQKQWAKRKSECINQNPAAASNHWQIPELARKVNTQKHTEFPQPVFSVSKQSPPI	423
DB	131	-----SNANPHAVSQGNLFSCC-----PSSKVSNIILHPNKDATASEMPA	170
QY	424	STSKWFDPKSICKTPSSNTLDYNSCFRTPVVKNDFFPACOLSTPYQP-----ACFOQQQ	479
DB	171	STN---DPEVRVKETDTSKQQQITGLEAPVG-----SSIYSGDGOANARLPEEL	217
QY	480	HQILATPLQNL---QVLASSANEC-----ISVKGRIYSILKQIGSGGSSKVFO	525
DB	218	HTSVSSQPKSKDKEKVGASGPAPKRNVDPLFPFKVNGKLYQRLGKIGSGGSSVHK	277
QY	526	VLNEKKQIYAIKYVNLBEADNQTLDSYRNEIAYLNKLQHSDDKIIRLYDYEITDQ----	580

DB	278	VISSDCTIYALKKIKLKGSDYATAYGFCQEIYGLKKLKGTN-IIQLIDYEVTDKTLQE	336
QY	581	-----YIYVMCEGNTIDNSWLKKK-----KSIDPWERKSYWKNM	615
DB	337	VLNGTMSNKGDKGVKEDGFYVWLEYGEIDLAHMLSKQWREIEGSDRTTIDENWLFYWOQI	396
QY	616	LEAVHTIHQHGIVHSDLPANFLIYDGMKLIDFGIANQMOPDTSVVVKDSQVGTWYMP	675
DB	397	LOAVNTIHEERIVHSDLPANFLVGRFELIDFGIAKAINSDTTNIQRDSQVGLSYMS	456
QY	676	PEAIKDMSSSRENGSKSKISPKSDVWSLGCILYTYWYKTFPQQIINQISKLHAIIDPN	735
DB	457	PEAFM-CNESDENGT-IKCGRPSDIWSLGCILYQWYVGRTPFPADYKTFWAKFKVITDPN	514
QY	736	HEIEFPDIPKDLQDLVLCCKLRDPKORISIPELLAHEYV	775
DB	515	HEITYNQLSNPWLIDIMKKCLANDRNQRWRIPELLQHPFL	554

Search completed: September 29, 2004, 16:16:14
Job time: 141 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 16:10:15 ; Search time 34 Seconds
(without alignments)
1276.984 Million cell updates/sec

Title: US-10-081-119-14

Perfect score: 4376
Sequence: 1 MNKVRDIKNKFKNEDLTDEL.....GGESHNSSSKTFEKKRGKK 841

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCITUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	310.5	7.1	890	1 US-08-472-934-8	Sequence 8, Appli
2	310.5	7.1	890	2 US-08-323-460A-8	Sequence 8, Appli
3	310.5	7.1	890	3 US-08-461-146C-8	Sequence 8, Appli
4	310.5	7.1	890	4 US-08-461-145C-8	Sequence 12, Appl
5	310.5	7.1	890	5 US-08-628-829-12	Sequence 4, Appli
6	306	7.0	505	6 US-09-729-995-4	Sequence 4, Appli
7	306	7.0	505	7 US-10-135-689-4	Sequence 10, Appl
8	303.5	6.9	1247	1 US-08-472-934-10	Sequence 10, Appl
9	303.5	6.9	1247	2 US-08-323-460A-10	Sequence 10, Appl
10	303.5	6.9	1247	3 US-08-461-146C-10	Sequence 10, Appl
11	303.5	6.9	1247	4 US-08-461-145C-10	Sequence 13, Appl
12	303.5	6.9	1597	4 US-09-423-890-13	Sequence 14, Appl
13	303.5	6.9	1597	4 US-08-628-829-14	Sequence 2, Appli
14	298.5	6.8	505	4 US-09-729-995-2	Sequence 2, Appli
15	298.5	6.8	505	4 US-10-135-689-2	Sequence 23, Appl
16	294	6.7	279	4 US-08-739-875-23	Sequence 2, Appli
17	291	6.6	416	1 US-08-232-995D-2	Sequence 2, Appli
18	291	6.6	416	2 US-08-834-108-2	Sequence 2, Appli
19	291	6.6	464	1 US-08-252-995D-6	Sequence 6, Appli
20	291	6.6	464	2 US-08-834-108-6	Sequence 6, Appli
21	291	6.6	647	3 US-09-031-563-7	Sequence 7, Appli
22	291	6.6	647	4 US-09-392-277-7	Sequence 7, Appli
23	291	6.6	647	4 US-09-258-000-7	Sequence 7, Appli
24	291	6.6	648	3 US-09-031-563-5	Sequence 5, Appli
25	291	6.6	648	4 US-09-392-277-5	Sequence 5, Appli
26	291	6.6	648	4 US-09-258-000-5	Sequence 5, Appli
27	291	6.6	925	1 US-08-252-995D-4	Sequence 4, Appli

28 291 6.6 925 2 US-08-834-108-4 Sequence 4, Appli
29 291 6.6 1315 3 US-08-031-563-2 Sequence 2, Appli
30 291 6.6 1315 4 US-09-293-505-10 Sequence 10, Appli
31 291 6.6 1315 4 US-09-392-277-2 Sequence 2, Appli
32 291 6.6 1315 4 US-09-258-000-2 Sequence 2, Appli
33 290 6.6 260 3 US-09-031-563-24 Sequence 24, Appli
34 290 6.6 260 4 US-09-392-277-24 Sequence 24, Appli
35 290 6.6 260 4 US-09-258-000-24 Sequence 24, Appli
36 289.5 6.6 544 4 US-08-688-188B-95 Sequence 95, Appli
37 289.5 6.6 544 4 US-09-291-417D-95 Sequence 95, Appli
38 288 6.6 1315 3 US-09-031-563-25 Sequence 25, Appli
39 288 6.6 1315 4 US-09-392-277-25 Sequence 25, Appli
40 288 6.6 1315 4 US-09-258-000-25 Sequence 25, Appli
41 285 6.5 544 3 US-08-559-397A-29 Sequence 29, Appli
42 284 6.5 274 2 US-07-857-224B-29 Sequence 29, Appli
43 283.5 6.5 275 4 US-09-803-671B-5 Sequence 5, Appli
44 282.5 6.5 431 2 US-09-211-930-5 Sequence 5, Appli
45 282.5 6.5 431 3 US-09-340-993-5 Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-472-934-8
; Sequence 8, Application US/08472934
; Patent No. 5753446
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL
; TITLE OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive and Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/472,934
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/440,421
; FILING DATE: 15-May-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/354,516
; FILING DATE: 21-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 5,405,941
; FILING DATE: 15-Apr-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/323,460
; FILING DATE: 14-Oct-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/11690
; FILING DATE: 14-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04178
; FILING DATE: 15-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr., Esq.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: CPI-004DVCP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941

QY 264 NS----PDCVKTDDSVVPCFMKQTSRSECDRLVVPKSGNDSCELRNLKSV----- 314
Db 217 NAATGKDCSKDPDDVFMDFALLTKHGDRARD-----SDGWTWEARAVKIVPQVETV 270
QY 315 -----QNSHFKEPLVDSKSELIIITDITLKNKTESS--LLAK-LEET 355
Db 271 DTLRSMQVDNLLLVVMEASAH-----LVLRKAFQOSIEGLMTVRHEQTSSQPIAKGLQOL 326
QY 356 K-----EYQEPVPSNQKQWAKRSEKINQNPAASSNHQ 392
Db 327 KNDALCNRISDAIDRVDMFTLEF--DAEVESESATLQOYYREAMIQYNGFGEYHKE 385
QY 393 I-----PELARKVNTQKHTEQVFSV-----SKQSPPISTSKWFD-----PKSI 434
Db 386 VVRLMSGEPRQKIG--DKYITFAQKXNYVLTKCESGRGTRPRWATQGFDFLQAIAPAFI 443
QY 435 CKTPSSNTLDDYMS-----CFRTPVKNDFPPAC---QLSTPYGQAPCFQO--QQHOI 482
Db 444 SALPE-----DDFLSLQALMNECIGHVIGKPHSPVTAIHRNSPRPVKPRCHSDPPNPHLI 499
QY 483 LATP-----LO----- 490
Db 500 IPTPEGRGSGVPENDRLASTAELQPSLRHSPTTEERDEPAYPRSDSGSTRRSWEL 559
QY 491 QVLASSANECISVKRIYSILK----- 513
Db 560 RTLLISQT--KDSASQGPALQKSVRLFERRRYEMRRKNIIQGVCDTPKSYDNVMHVL 618
QY 514 -----QISGSSKVFQVLN--EKQIVAKVNVLEADNQTLDSYRNEIAYLNK 561
Db 619 RKTFTKQWGNKIGEGYKVTTCISVDTGELMAMKEIRFPNDHKTIKETABELKIFEG 678
QY 562 LQORSDKIIRLYDEITDQIYMWME--C--GNIDLSNWLKKSIDSIPWERSKYWKMLEA 618
Db 679 IKH--PNLVRYFGVELHREMYIFWECDEGLTLEVSRLGQEHV-----IRLYTKQITVA 732
QY 619 VHTHQHGVHSDUKPAN--FLIVDGMKLKIDFGIANOMQDPTTSVVD--SQVGTVNYMP 675
Db 733 INVLRHGVHVDIKGANIELTSSGLIKLGDGFGSVKLKKNNAOTMPCGVENVSTLTAAVYA 792
QY 676 PEATKDNSSRENGSKSPKSDVMSLGLCYLYMTYTGKTPFOCIINQISKLHAIIDPN 735
Db 793 PEVI-----THAKGEHGR---AADIWSLGVVEMVTKRPHWEYEHNFQIMYK--VGMG 843
QY 736 HEIEFPDIPEX---DLQDVLKCCLRDPPKQKISPELAPHYVQIQT 779
Db 844 HK---PIPERLSPEGRAFLSHCLESDDPKIRWTASQLLDHAFVKVCT 887

RESULT 3
US-08-461-146C-8
; Sequence 8, Application US/08461146C
; Patent No. 5981265
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: METHODS FOR REGULATING MEKK PROTEIN ACTIVITY
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive and Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,146C
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/354,516
; FILING DATE: 21-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/049,254
; FILING DATE: 15-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/323,460
; FILING DATE: 14-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/11690
; FILING DATE: 14-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04178
; FILING DATE: 15-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: KARA, Catherine J.
; REGISTRATION NUMBER: P41,106
; REFERENCE/DOCKET NUMBER: CPI-004CN3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 890 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Xaa = Any amino acid
; LOCATION: 116
; US-08-461-146C-8

Query Match 7.1%; Score 310.5; DB 2; Length 890;
Best Local Similarity 20.2%; Pred. No. 2.9e-16;
Matches 191; Conservative 150; Mismatches 307; Indels 299; Gaps 42;

QY 39 IMMANNPDDWLSLLKLEKNSVPLSDAI-----LNKILGRYSQAIEA-----LPDPKYG- 88
Db 34 LMVFDYRSYTMQLQQLPQASHSLKNLLEEWNFTEITHYIRGEAQAGKLFCDIAGM 93
QY 89 ---QNESEF--ARIOVRFAELKAIQPPDDARDYFQVARANCKKFAFVHISFAQFELSQGNV 143
Db 94 LKSTGSLFESGLQESCAELWTXADDNGAAD--ELRR-----SVIEIS----- 134
QY 144 KKSQQLLQKAVRGAVPLEMLEIALNLNLKQKLLSEEEKNLKSASTVLTQESFSGSL 203
Db 135 RALKELFHEAREASKALGFAMLRKDLKLEIAAEFVLSASARELLDA--LKAKQYVKVQI 191
QY 204 GHLONRNNSCDSEGGQTTKARFLYGENMPPQDAEIGYRNSLRQTNKTKQSCPPGRYPVNL 263
Db 192 PGLN-----LHVFPDLSAEKK-----ILQL 216
QY 264 NS---PDCVKTDDSVVPCFMKQTSRSECDRLVVPKSGNDSCELRNLKSV----- 314
Db 217 NAATGKDCSKDPDDVFMDFALLTKHGDRARD-----SDGWTWEARAVKIVPQVETV 270
QY 315 -----QNSHFKEPLVDSKSELIIITDITLKNKTESS--LLAK-LEET 355
Db 271 DTLRSMQVDNLLLVVMEASAH-----LVLRKAFQOSIEGLMTVRHEQTSSQPIAKGLQOL 326
QY 356 K-----EYQEPVPSNQKQWAKRSEKINQNPAASSNHQ 392
Db 327 KNDALCNRISDAIDRVDMFTLEF--DAEVESESATLQOYYREAMIQYNGFGEYHKE 385
QY 393 I-----PELARKVNTQKHTEQVFSV-----SKQSPPISTSKWFD-----PKSI 434
Db 386 VVRLMSGEPRQKIG--DKYITFAQKXNYVLTKCESGRGTRPRWATQGFDFLQAIAPAFI 443
QY 435 CKTPSSNTLDDYMS-----CFRTPVKNDFPPAC---QLSTPYGQAPCFQO--QQHOI 482
Db 444 SALPE-----DDFLSLQALMNECIGHVIGKPHSPVTAIHRNSPRPVKPRCHSDPPNPHLI 499

QY 483 LATP-----LQ-----NL 490
 Db 500 IPTPEGRGSSVPENDRLASIAELQPSLSRHSSTPEERDEPAYPRSDSGSTRSWEL 559
 QY 491 QVLASSANECSVKGRYISLK-----513
 Db 560 RTLSQ-TKDSASQGPPIEAIQSVRLFEERRVREMRKNIIGQVCDTPKSYDNVHVGL 618
 QY 514 -----QIGSGSSKVFQVLN-EKKQIYAIKYNLEADNQTLDYSRNEIAYLNK 561
 Db 619 RVTFFKQGNKIGEGYKVTCTISVDTGELMAMKEIRFQPNHKTIKETABELKIFEG 678
 QY 562 LQHSXKIIRLYDIETDQIYVWME-C--GNIDLSNWLKKSIDPWERKSYWKNLEA 618
 Db 679 IKH--PNLVRYFGVELHREEMYPMEYCDGTLEESVRLGQEHV----IRLYTKQITVA 732
 QY 619 VHTHQHGIHSDILKPN-FLIVDGMKLIDFGIANOMQDPTTSVKD--SQVGTVNYMP 675
 Db 733 INVLEHGIHVRDITKGANIFLTSSGLIKLDFGCSVKLNNAQTMPEGVNSTLGTAYMA 792
 QY 676 PEAIKDMSSRENGKSKSPKSDVMSLGLIYVYTYGTPQOIIINOISKLHAIIDPN 735
 Db 793 PEVI-----TRAKGEHGR--ADIWSLGCVIEVWTKRPHWEYEHFQIMYK-VGMG 843
 QY 736 HEIEFPDPEK--DLQDVLLKCLKRDPKQKORISIPELLAHPYVQIQT 779
 Db 844 HK--PPPERLSPEGRKFLSHCLSDPKIRWTASQLLDHAFVKVCT 887

RESULT 4

US-08-461-145C-8
 ; Sequence 8, Application US/08461145C
 ; Patent No. 6074861
 ; GENERAL INFORMATION:
 ; APPLICANT: JOHNSON, GARY L.
 ; TITLE OF INVENTION: NOVEL MEKK PROTEINS
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lathive and Cockfield, LLP
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/461,145C
 ; FILING DATE: 5-JUNE-1995
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/049,254
 ; FILING DATE: 11-APR-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/323,460
 ; FILING DATE: 14-OCT-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US94/11690
 ; FILING DATE: 14-OCT-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US94/04178
 ; FILING DATE: 15-APR-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/354,516
 ; FILING DATE: 21-FEB-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kara, Catherine J.
 ; REGISTRATION NUMBER: P-41,106
 ; REFERENCE/DOCKET NUMBER: CFI-004CN1
 ; TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 890 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; FEATURE: Xaa = Any amino acid
 ; LOCATION: 116
 ; US-08-461-145C-8

Query Match 7.1%; Score 310.5; DB 3; Length 890;
 Best Local Similarity 20.2%; Pred No. 2.9e-16;
 Matches 191; Conservative 150; Mismatches 307; Indels 299; Gaps 42;

QY 39 IMMANNPEDWLSLLKLEKNSVPLSDAL-----LNKLIGRYSOAIEA---LPPDKYG- 88
 Db 34 LMVYFDYMRSMIOMLQQLPOASHSLKNLLEBEMNFTKEITHYIRGGEAQAGKLFCDIAGM 93
 QY 89 ---QNESF--ARIQVRAELKAIQEPDARDYFOVARANCKKFAFVHISFAQFELSQGVN 143
 Db 94 LLKSTGSLFESGLQESCAELMTXADDNGAAD--ELRR-----SVIEIS----- 134
 QY 144 KSKQLLQKAVERGAVPLEMLEIALRNINLQKKQLLSEEEKKNLSASTVLTQAESFSGSL 203
 Db 135 RALKELPHEAERASKALGFAKMLRKOLEIAAEFVLSASARELLDA--LKAQYVVKVQI 191
 QY 204 GHLQNRNNSCDRSGTQTKAREFLYGENMPPQDAEIGYRNSLRQNTKTKQSCFFGRVFNLL 263
 Db 192 PGLN-----LHVFPDLSAEKK-----IILQLL 216
 QY 264 NS---PDCDVKTDBSVVPCFMKQRTSRSECRDLVVPGSKPSGNDSCBLNLSKV----- 314
 Db 217 NAATGKCCSKDPDDVFDADFLLTKHGDRARD-----SEDGWTWEARAVKIVPQVETV 270
 QY 315 -----QNSHFKEPLVDSKXSELIITDSITLKNKTS--LLAK-LEET 355
 Db 271 DTLRSMQVDNLLLVVMESAH-----LVLRKAFQOSIEGLMTVRHEQTSSQPIIAKGLQQL 326
 QY 356 K-----EYQBPVPSNQKQWQAKRSECEINQNPAASSNHQ 392
 Db 327 KNDALCNRISDAIDRVDEMFTLEF--DAEVESESATLQYREAMIQYNGFGEFHK 385
 QY 393 I-----PELARKVNTQKHITFEQPFVS-----SKQSPPISTSKWFD-----PKSI 434
 Db 386 VVRLMSGEFRQKIG--DKYITFAKWMYVLTCKESGRGTRPRWATQGFLOAIEPAFI 443
 QY 435 CKTPSSNTLDDYMS-----CFRTPVVKNDFPAC--QLSTPYGQAPACFQ--QOHOI 482
 Db 444 SALPE---DDFLSLQALMNECIGHVIGKPHSPVTAIHRNSPRVKVPRCHSDPPNPHLI 499
 QY 483 LATP-----LQ-----NL 490
 Db 500 IPTPEGRGSSVPENDRLASIAELQPSLSRHSSTPEERDEPAYPRSDSGSTRSWEL 559
 QY 491 QVLASSANECSVKGRYISLK-----513
 Db 560 RTLSQ-TKDSASQGPPIEAIQSVRLFEERRVREMRKNIIGQVCDTPKSYDNVHVGL 618
 QY 514 -----QIGSGSSKVFQVLN-EKKQIYAIKYNLEADNQTLDYSRNEIAYLNK 561
 Db 619 RVTFFKQGNKIGEGYKVTCTISVDTGELMAMKEIRFQPNHKTIKETABELKIFEG 678
 QY 562 LQHSXKIIRLYDIETDQIYVWME-C--GNIDLSNWLKKSIDPWERKSYWKNLEA 618
 Db 679 IKH--PNLVRYFGVELHREEMYPMEYCDGTLEESVRLGQEHV----IRLYTKQITVA 732
 QY 619 VHTHQHGIHSDILKPN-FLIVDGMKLIDFGIANOMQDPTTSVKD--SQVGTVNYMP 675
 Db 733 INVLEHGIHVRDITKGANIFLTSSGLIKLDFGCSVKLNNAQTMPEGVNSTLGTAYMA 792

```
QY 676 PEAIKDMSSRENGSKSPKSDVMSLGCILYMTYGTPTFQIINOISKLHAIIDPN 735
DB 793 PEVI-----TRAKGEHGR---AADIWSLGCVVIEWTGRPMHEYEHNFQIMYK-VGMG 843
QY 736 HEIEFPDIPEK---DLQDVLMKCLKRPKORISIPPELLAHPIYVQIQT 779
DB 844 HK---PPIPERLSPEGKAFLSHCLESDEPKIRWTASQLLDHAFVKVCT 887

RESULT 5
US-08-628-829-12
; Sequence 12, Application US/08628829A
; Patent No. 6333170
; GENERAL INFORMATION:
; APPLICANT: Johnson, Gary L.
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To External
; FILE REFERENCE: CFI-004DVC3
; CURRENT APPLICATION NUMBER: US/08/628,829A
; EARLIER FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: 08/410,421
; EARLIER FILING DATE: 1995-05-15
; EARLIER APPLICATION NUMBER: 08/323,460
; EARLIER FILING DATE: 1994-10-14
; EARLIER APPLICATION NUMBER: 08/049,254
; EARLIER FILING DATE: 1993-05-15
; EARLIER APPLICATION NUMBER: 08/410,602
; EARLIER FILING DATE: 1995-04-24
; EARLIER APPLICATION NUMBER: 08/472,934
; EARLIER FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 12
; LENGTH: 890
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: At position 116, Xaa = any amino
; OTHER INFORMATION: acid
US-08-628-829-12

Query Match 7.1%; Score 310.5; DB 4; Length 890;
Best Local Similarity 20.2%; Pred. No. 2.9e-16;
Matches 191; Conservative 150; Mismatches 307; Indels 299; Gaps 42;

QY 39 IMMANNPEDWLSLLKLNKSNVPSDAL-----LNKLGRYSQAIEA-----LPPKYG- 88
DB 34 LMVFDYMRSWIQMLQLOLPQASHSLKNLLEBWNFTKEITHYIRGGEAQAGKLFCDIAGM 93
QY 89 ---QNESEF--ARIVRAELKALICEPDDARDYFOMARANCKFAFVHISPAQFELSQGNV 143
DB 94 LKSTGSLGSLQESCAELWTXADDNGAAD--ELRR-----SVIEIS----- 134
QY 144 KKSQQLQKAVERGAVPLEMLEIALRNINLQKKQLLSEEEKKNLSASTVLTAQESFSGSL 203
DB 135 RALKELPHEARERASKALGPAKMLRKDLTAAPVLSASARELLDA--LKAQYVKVQI 191
QY 204 GHLQNRNNSCDSRQTTKARFLYGENMPPQDAEIGYNSLRQNTKTKQSCPPGRVPVNL 263
DB 192 PGLN-----LHVFPDPSLABEKK-----IILQLL 216
QY 264 NS-----PDCDKVTDVSVPCFMKQTSRSECRDLVPGSKPSGNDSCBLRLKSV----- 314
DB 217 NAATGDKCDKDPDDVDFDAFLLTKHGDRARD-----SEDGWTWEARAVKIVPQVETV 270
QY 315 -----QNSHPKEPLVSEKSSSELIITDSTLKNKTESS--LIAK-LEET 355
DB 271 DTLASMQVDNLLLVMMESAH---LVLQKRAFOQSIEGLMTRVHEQTSOPIIAKGLQQL 326
QY 356 K-----EYQEPPEPNQKQWAKRKSECTINQNPAASSHMQ 392
DB 327 KNDALELNCRISDAIRDVHEMFTLEF--DAEVESESATLQOYREAMIQYNGFEYHKE 385
QY 393 I-----PELARKVNTQKHTTFEQVFVS-----SKQSPPISTSKWFD-----PKSI 434
```

```
DB 386 VRLMSGEPFRQIG---DKYITFAQKMMNVLTKCSGRGTRPRWATQGFDFQALIEPAFI 443
QY 435 CKTPSNTLDDYMS-----CFRTPVVKNDPPPPAC---QLSTPYGQACFOQ--QQHQI 482
DB 444 SALPE-----DDELSLQALMNECIGHVIGKPSPTVAIHNSPRPVKVPRCHSDPENPHLI 499
QY 483 LATP-----LQ-----NL 490
DB 500 IPTPEGFGSSVPENDRLASIAAELOFRSLRSHSPTTERDEPAYPRSDSGSTRRWEL 559
QY 491 QVLASSANECISVKGRIYSILK-----VRLFEERYREMRKNIIGQVCDTPKSYDNVMHVGL 618
DB 560 RTLISQT-KDSASKQGPTEAIOKSVRLFEERYREMRKNIIGQVCDTPKSYDNVMHVGL 618
QY 514 -----QIGSGGSKVFOVLN-EKKQIYAIKYVNLBEADNQTLDVSRNEIAYLNK 561
DB 619 RKVTFKQWGNKIGEGQYKVTCSVDTGELMAMKEIRFQENDHKTKETADELKIPEG 678
QY 562 LOOHSDKIIRLYDEITQYIYMME-C--GNIDLSWLKKKSDPNERKSYWQMLEA 618
DB 679 IKH--PNLVRYFVGLHREEMVIFMEYCDEGTLEEVSRGLQEHV-----IRLYTKQITVA 732
QY 619 VHTIHOHGIVHSDLPKAN-FLIVDGLMLKIDFGIANQMPDPTTSVVKD--SOVGTVNYMP 675
DB 733 INVLEHGIVHRDIKGANIFLTSSGLIKLGFQCSVKLKNNAQTMPGEVNSTLGTAAAYMA 792
QY 676 PEAIKDMSSRENGSKSPKSDVMSLGCILYMTYGTPTFQIINOISKLHAIIDPN 735
DB 793 PEVI-----TRAKGEHGR---AADIWSLGCVVIEWTGRPMHEYEHNFQIMYK-VGMG 843
QY 736 HEIEFPDIPEK---DLQDVLMKCLKRPKORISIPPELLAHPIYVQIQT 779
DB 844 HK---PPIPERLSPEGKAFLSHCLESDEPKIRWTASQLLDHAFVKVCT 887

RESULT 6
US-09-729-995-4
; Sequence 4, Application US/09729995
; Patent No. 6426206
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO00904
; CURRENT APPLICATION NUMBER: US/09/729,995
; CURRENT FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-729-995-4

Query Match 7.0%; Score 306; DB 4; Length 505;
Best Local Similarity 25.1%; Pred. No. 2.9e-16;
Matches 117; Conservative 73; Mismatches 155; Indels 122; Gaps 21;

QY 349 LAKLEETKEYQPEVPESNQKQWAKRKSECTINQNPAASSHMQIPELARKVNTQKHTT 408
DB 25 VAHLEAEAGEP-----ASNGVDPPPARAASVIPGSAS 59
QY 409 FQQPVFSVKOSPPTTSKWFDPKSICTKTPSNTLDDVMSCFRTPVVKNDFPPACQLSTP 468
DB 60 RPTPV-----RPSLSARKF---SLQERPAG-----SCLEAQVGFSYTGTPASHMS-- 100
QY 469 YQCPACFOQ---QQHQILATPLQNLQVLASSANECISVKGRIVSILKQIGSGSSKVFQ 525
DB 101 ----PRAWRRPTIESHHV-----AISTEDCVQLNQ--YKLOSEIGKRGAYGVVRL 144
QY 526 VLNEKK-QIYAIKYVNLBEADNQ-----TLDYSYRNEIAYLN 560
```

Db 145 AYNEREDRHVAMKVLKXKLLKQYGFPRPPRGSQAPQGGPAKQQLPLERVVQETAILK 204
Qy 561 KLOQ-HSDKIRLYDYITQYIYMWMECGNIDLNSWLKX-----KSIDPWERKSY 611
Db 205 KLDHNVVVKLIEVLDDPAEDN-LYLVPDL-----LRKGPVMEVPCDKPFPEEQARLY 255
Qy 612 WKWLEAVHTTHQGIHVHSDLKANFLI-VDMKMLKIDFGIANOMQPDITTSVVYKDSQVGT 670
Db 256 LRDIILGLEVLHCOKIIVHRDIKPSNLLGDDGHVKIADFGVSNQFEGNDAQL--SSTAGT 313
Qy 671 VNYMPPAIDMSSSRRENGSKSKISPKSDVWSLGCILYMTYTKTPFOOINQISKLHA 730
Db 314 PAFVAPPAISDTGOS-FSGKA-----LDVWATGVTLVCFVYKGCFF--IDEVILALHR 363
Qy 731 IIDPNHEIEFPDIE--KLDQVLKCLKRDPKORISIPPELLAHYV 775
Db 364 KI-KNEAVVPPEEVESEELKDLILKMLDKNPETRIGVSDIKLHPW 409

RESULT 7
US-10-135-689-4
; Sequence 4, Application US/10135689
; Patent No. 6670162
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00904DIV
; CURRENT APPLICATION NUMBER: US/10/135,689
; CURRENT FILING DATE: 2002-05-01
; PRIOR FILING DATE: 60/247,031
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: 09/729,995
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-135-689-4

Query Match 7.0%; Score 306; DB 4; Length 505;
Best Local Similarity 25.1%; Pred. No. 2.9e-16;
Matches 117; Conservative 73; Mismatches 155; Indels 122; Gaps 21;
Qy 349 LAKLEETKEVQEPVPSNQKQAKKSECIQNQPAASNNHQIPELARKVNTQKHTT 408
Db 25 VAHLEAEAEQEP-----ASGVDPPPRAAASVIEGSA 59
Qy 409 PEQVFSVSKOSPPISKMPDKSIKCTPSSNTLDDYMSCFRTPVVQNDFFPACQLSTP 468
Db 60 RPTPV-----RPSLSARKE---SLQERPAQ-----SCLEAQVGPYSTGPAHMS-- 100
Qy 469 YGQACFQQ---QHQILATPLQNLQVLASSANECSIVKGRISYILKQIGSGSGSKVQ 525
Db 101 ---PRAWRRPTIESHHV-----AISTEDCVQLNQ--YKQSEIGKAGYGVWL 144
Qy 526 VLNEKK-QIYAIVKYNLEADNQ-----TLDYSYNEIAYLN 560
Db 145 AYNEREDRHVAMKVLKXKLLKQYGFPRPPRGSQAPQGGPAKQQLPLERVVQETAILK 204
Qy 561 KLOQ-HSDKIRLYDYITQYIYMWMECGNIDLNSWLKX-----KSIDPWERKSY 611
Db 205 KLDHNVVVKLIEVLDDPAEDN-LYLVPDL-----LRKGPVMEVPCDKPFPEEQARLY 255
Qy 612 WKWLEAVHTTHQGIHVHSDLKANFLI-VDMKMLKIDFGIANOMQPDITTSVVYKDSQVGT 670
Db 256 LRDIILGLEVLHCOKIIVHRDIKPSNLLGDDGHVKIADFGVSNQFEGNDAQL--SSTAGT 313
Qy 671 VNYMPPAIDMSSSRRENGSKSKISPKSDVWSLGCILYMTYTKTPFOOINQISKLHA 730

Db 314 PAFVAPPAISDTGOS-FSGKA-----LDVWATGVTLVCFVYKGCFF--IDEVILALHR 363
Qy 731 IIDPNHEIEFPDIE--KLDQVLKCLKRDPKORISIPPELLAHYV 775
Db 364 KI-KNEAVVPPEEVESEELKDLILKMLDKNPETRIGVSDIKLHPW 409

RESULT 8
US-08-472-934-10
; Sequence 10, Application US/08472934
; Patent No. 5753446
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL
; TITLE OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive and Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/472,934
; APPLICATION NUMBER: US/08/472,934
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/440,421
; FILING DATE: 15-May-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/354,516
; FILING DATE: 21-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 5,405,941
; FILING DATE: 15-Apr-1993
; APPLICATION DATA:
; APPLICATION NUMBER: 08/323,460
; FILING DATE: 14-Oct-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/11690
; FILING DATE: 14-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04178
; FILING DATE: 15-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr. Esq.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: CPI-004DVCP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1247 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-472-934-10

Query Match 6.9%; Score 303.5; DB 1; Length 1247;
Best Local Similarity 22.2%; Pred. No. 1.7e-15;
Matches 113; Conservative 99; Mismatches 165; Indels 131; Gaps 23;
Qy 296 PGSKPSGNDSCELNLSKVNQSHFKEPLVSKSESLIITDSITLKNKTSLLAKLE-- 353
Db 844 PG--PGGSDSVPAKPVNT-----APDTRGSSVPEMDRL-----ASIAAELOQR 884

QY	354	-----ETKEYQPEVPESN-----QKQWQAKRKSECCINQNPAAASNHHWQIPELARKVN	401
Db	885	SLSRHSPTERDEPAYPRSDSGSTRSWELR--TLISQKDSASKQGPBAIQKSVR	941
QY	402	TEQKHTTFEPQVSVSKQSPPISTKSWFDPKSICTPSSNTLDDYMSCFRTPVVKNDPPP	461
Db	942	-----LFEERRYREMRKRNIG-----QVCDTPKS--YDNVHVHGLURKVT-----	979
QY	462	ACQLSTPYQAPAFQOQOQHOILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSS	521
Db	980	-----FKWQRG-----NKIGEQYG	994
QY	522	KVFQVLN-EKKQIYAIKYNLEADNQTLDYSRYNETIAYLNKLOHSDKIIRLDYDEITDQ	580
Db	995	KVYTCISVDVTELMAMKEIFQPNDHKTIKETADELXIFEGIKH--PNLVRYFGVELHRE	1052
QY	581	YIYVMWE-C-GNIDLNSWLKKKSIDPWERKSVKMNLEAVHTIHOHGIVHSDLKPAN-	636
Db	1053	EMVIFMEYCEGTLEEVSRGLQEHV---IRLYTKQITVAINVLHEHGIVHREDIKCANI	1108
QY	637	FLIVDGMKLIDFGIANQMPDTSVVKD--SQVGTVYNNMPPEAIKDMSSSRNGSKSKS	694
Db	1109	FLTSSGLIKLGDGFCSCVKKNNATQMEGVNSTLGTAAVMAPEVI-----TRAKGEHGHR	1163
QY	695	ISPKSDWSJGCGILYNYMTCYKTFQOILINQISKLHAIIDPNHRIERPDIPEK--DLODV	751
Db	1164	---AAIWSJGCVVIEWTCRPHWEYHNFQIMYK-VGMCHK---PPIPERLSPEGKAF	1216
QY	752	LKCLKRDPKQR-SIPELLAHYPVQIQT	779
Db	1217	LSHCLESDPKIRWTASQLLDHAFVKVCT	1244
RESULT 9			
US-08-323-460A-10			
; Sequence 10, Application US/08323460A			
; Patent No. 5854043			
; GENERAL INFORMATION:			
; APPLICANT: JOHNSON, GARY L.			
; TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL			
; TITLE OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS			
; NUMBER OF SEQUENCES: 10			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: SHERIDAN ROSS & MCINTOSH			
; STREET: 1700 LINCOLN STREET, SUITE 3500			
; CITY: DENVER			
; STATE: CO			
; COUNTRY: USA			
; ZIP: 80203			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/323,460A			
; FILING DATE: 14-OCT-1994			
; CLASSIFICATION: 530			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 08/049,254			
; FILING DATE: 14-APR-1993			
; ATTORNEY/AGENT INFORMATION:			
; NAME: KOVARIK, JOSEPH E.			
; REGISTRATION NUMBER: 33,005			
; REFERENCE/DOCKET NUMBER: 2879-1-1			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 303/863-9700			
; TELEFAX: 303/863-0223			
; INFORMATION FOR SEQ ID NO: 10:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1247 amino acids			
; TYPE: amino acid			

APPLICATION NUMBER: US 08/049,254
 FILING DATE: 15-APR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/323,460
 FILING DATE: 14-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/11690
 FILING DATE: 14-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/04178
 FILING DATE: 15-APR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: KARA, Catherine J.
 REGISTRATION NUMBER: P41,106
 REFERENCE/DOCKET NUMBER: CPI-004CN3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1247 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-461-146C-10

Query Match 6.9%; Score 303.5; DB 2; Length 1247;
 Best Local Similarity 22.2%; Pred. No. 1.7e-15;
 Matches 113; Conservative 99; Mismatches 165; Indels 131; Gaps 23;

QY	296	PGSPGNDSC	ELNKLKSVQNSHFKPLVDSKSSLLITD	ITLKNKTESLLAKLE--	353
DB	844	PG--PGGDSV	PAKPVNT-----	ADTRGSSVPENDRL-----	ASIAAELOQR 884
QY	354	-----	ETKEYQPEVPESN-----	QKQWAKKSECNQNPAA	SNHWHQIPELARKYN 401
DB	885	SLSRHSSPT	ERDEPAYPRSDSGSTRSWE	L---TLISQKDSASKQGP	IEAIQKSVR 941
QY	402	TEQKHTT	FEQPVFSVSKSPIS	TSKWFDPKSI	CKTPSSNTLDDYMSCFRTPVVKNDFPP 461
DB	942	-----	LFERRRYEMRKNIIG-----	QVCDTPKS---YDVMHVG	LKRVT----- 979
QY	462	ACQLSTPY	GPACFQQQHQI	LATPLQNLQVLASS	ANECISVKGRYISILKQIGSGSS 521
DB	980	-----	FKWQRG-----	NKIGEGQYG 994	
QY	522	KVFQVLN-	EKKQIYAIKVN	LEADNQTLD	SYRNEIAYLNKLOQHSKIIRLDYETDQ 580
DB	995	KVYTCISV	DTGELMAKKEIRFQ	PNDHKTIKET	ADCLKIFEGIKH--PNLVRYFGVELHRE 1052
QY	581	YIYVWME-	C--GNIDLSN	WLKKKSIDP	WERKSYWKNMLEAVHTIHQHGIVHSDLPAN- 636
DB	1053	EMVIFMEY	CDGTELEVSRLQEHV---	IRLYTKQITVA	INVLEHGHVHRDIKGANI 1108
QY	637	FLIVDGM	LKIDFGIANQMPD	TSVVKD--	SQVGTVMYPPPAIKDMSSRENGSKSK 694
DB	1109	FLTSSGLI	KLGDFGCSVKLN	QAQMPGEVNST	LGTAAYNAPEVI-----TRAKGEHGR 1163
QY	695	ISPKSDV	MSGLIYMTYGTPT	QQQIINQISK	LHAIDPNHIEPDIPEK---DLQDV 751
DB	1164	---RAD	WLSGCVVMTGK	RPHEYEHFQ	IMYK-VGNGHK---PPIPERLSPGKAF 1216
QY	752	LKCKLRD	PKRISIP	PELLAH	PPYVQIQT 779
DB	1217	LSHCLESD	PKIRWTASQL	LLOHAF	VKVCT 1244

RESULT 11

US-08-461-145C-10
 Sequence 10, Application US/08461145C
 Patent No. 6074861
 GENERAL INFORMATION:
 APPLICANT: JOHNSON, GARY L.

TITLE OF INVENTION: NOVEL MEKK PROTEINS
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lahive and Cockfield, LLP
 STREET: 60 State Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/461,145C
 FILING DATE: 5-JUNE-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/049,254
 FILING DATE: 11-APR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/323,460
 FILING DATE: 14-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/11690
 FILING DATE: 14-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/04178
 FILING DATE: 15-APR-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/354,516
 FILING DATE: 21-FEB-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Kara, Catherine J.
 REGISTRATION NUMBER: P-41,106
 REFERENCE/DOCKET NUMBER: CPI-004CN1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1247 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-461-145C-10

Query Match 6.9%; Score 303.5; DB 3; Length 1247;
 Best Local Similarity 22.2%; Pred. No. 1.7e-15;
 Matches 113; Conservative 99; Mismatches 165; Indels 131; Gaps 23;

QY	296	PGSPGNDSC	ELNKLKSVQNSHFKPLVDSKSSLLITD	ITLKNKTESLLAKLE--	353
DB	844	PG--PGGDSV	PAKPVNT-----	ADTRGSSVPENDRL-----	ASIAAELOQR 884
QY	354	-----	ETKEYQPEVPESN-----	QKQWAKKSECNQNPAA	SNHWHQIPELARKYN 401
DB	885	SLSRHSSPT	ERDEPAYPRSDSGSTRSWE	L---TLISQKDSASKQGP	IEAIQKSVR 941
QY	402	TEQKHTT	FEQPVFSVSKSPIS	TSKWFDPKSI	CKTPSSNTLDDYMSCFRTPVVKNDFPP 461
DB	942	-----	LFERRRYEMRKNIIG-----	QVCDTPKS---YDVMHVG	LKRVT----- 979
QY	462	ACQLSTPY	GPACFQQQHQI	LATPLQNLQVLASS	ANECISVKGRYISILKQIGSGSS 521
DB	980	-----	FKWQRG-----	NKIGEGQYG 994	
QY	522	KVFQVLN-	EKKQIYAIKVN	LEADNQTLD	SYRNEIAYLNKLOQHSKIIRLDYETDQ 580
DB	995	KVYTCISV	DTGELMAKKEIRFQ	PNDHKTIKET	ADCLKIFEGIKH--PNLVRYFGVELHRE 1052
QY	581	YIYVWME-	C--GNIDLSN	WLKKKSIDP	WERKSYWKNMLEAVHTIHQHGIVHSDLPAN- 636

Db 1053 EMYIFMYCEGLLEVSRLGLQEHV-----IRLYTKQITVAINVLEHGHVIRHDIKGANI 1108
Qy 637 FLIVDGMKLIIDFGIANQMOPDTSVVKD--SQVGTVMYMPPEAIKDMSSRENGSKSK 694
Db 1109 FLTSSGLIKLGDGFCVSKLNNNAQTMEGVNSTLGTAAAYMAPEVI-----TRAKGEHGR 1163
Qy 695 ISPKSDVMSLGCILYNYTYGKTPFQOIINOISKLHAIIDPNHEIEPDIPEK---DLQDV 751
Db 1164 ---AADISLGCVVIVMTGKRPWHEYEHNFQIMYK-VGMGHK---PPPERLSPEGKAF 1216
Qy 752 LKCCLRDPKQKORISIPPELLAHYVQIQ 779
Db 1217 LSHCLESDPKIRWTASQLLDHAFVKVCT 1244

RESULT 12

US-09-423-890-13
; Sequence 13, Application US/09423890
; Patent No. 6312934
; GENERAL INFORMATION:
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: HUMAN MEKK PROTEIN AND NUCLEIC ACID MOLECULES
; FILE REFERENCE: CPI-085CPPC
; CURRENT APPLICATION NUMBER: US/09/423,890
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: USSN 60/078,153
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: USSN 60/099,165
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1597
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-423-890-13

Query Match 6.9%; Score 303.5; DB 4; Length 1597;
Best Local Similarity 22.2%; Pred. No. 2.5e-15;
Matches 113; Conservative 99; Mismatches 165; Indels 131; Gaps 23;
Qy 296 PGSKPSGNDSCBLNLSKVNSHFKPELVDSKSELIITDSTLKNKTESLLAKLE-- 353
Db 1194 PG--PGGGDSVPAPKPVNT-----APDTRGSSVPENDRL-----ASTAAELQFR 1234
Qy 354 -----ETKEYQPEVPESEN-----OKQWQAKRKSCINONPAASNNHWOIPELARKVN 401
Db 1235 SLRSHSPTERDEPAYPRSDSSSTRSWELR---TLISQTKDSASKQGPFAIQSVR 1291
Qy 402 TEQKHTTFEQPVFSVKOSPPISTSKWFDPKSICKTPSSNTLDDYMSCFTPVVKNDFPP 461
Db 1292 -----LFEERRRYEMRKNIIIG-----QVCDTPKS--YDVMVHVLGRKVT----- 1329
Qy 462 ACQLSTPYGPACFQOQHQILATPLQLVLASSANECISVKGRISYILKQISGGSS 521
Db 1330 -----FKWQRG-----NKIGEGQYG 1344
Qy 522 KVFOVLN-EKKQIYAIKVNLEADNQTLDPSYRNEIAYLNKLOQHSDDKIIRLYDYBITDQ 580
Db 1345 KVTTCISVDTGELMAKKEIRFPNDHKTIKETADELKIFEGIKH--PNLVRYFGVELHRE 1402
Qy 581 YIYVME-C--GNIDLSNLKXKSIDPWERKSYWKNWLEAVHTIHCHGIVHSDLKPAN- 636
Db 1403 EMYIFMYCEGTLLEVSRLGLQEHV-----IRLYTKQITVAINVLEHGHVIRHDIKGANI 1458
Qy 637 FLIVDGMKLIIDFGIANQMOPDTSVVKD--SQVGTVMYMPPEAIKDMSSRENGSKSK 694
Db 1459 FLTSSGLIKLGDGFCVSKLNNNAQTMEGVNSTLGTAAAYMAPEVI-----TRAKGEHGR 1513
Qy 695 ISPKSDVMSLGCILYNYTYGKTPFQOIINOISKLHAIIDPNHEIEPDIPEK---DLQDV 751
Db 1513 -----AADISLGCVVIVMTGKRPWHEYEHNFQIMYK-VGMGHK---PPPERLSPEGKAF 1566
Qy 695 ISPKSDVMSLGCILYNYTYGKTPFQOIINOISKLHAIIDPNHEIEPDIPEK---DLQDV 751

Db 1514 ---AADISLGCVVIVMTGKRPWHEYEHNFQIMYK-VGMGHK---PPPERLSPEGKAF 1566
Qy 752 LKCCLRDPKQKORISIPPELLAHYVQIQ 779
Db 1567 LSHCLESDPKIRWTASQLLDHAFVKVCT 1594
RESULT 13
US-0628-829-14
; Sequence 14, Application US/08628829A
; Patent No. 6333170
; GENERAL INFORMATION:
; APPLICANT: Johnson, Gary L.
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To External
; FILE REFERENCE: CPI-004DYCP3
; CURRENT APPLICATION NUMBER: US/08/628,829A
; CURRENT FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: 08/440,421
; EARLIER FILING DATE: 1995-05-15
; EARLIER APPLICATION NUMBER: 08/323,460
; EARLIER FILING DATE: 1994-10-14
; EARLIER APPLICATION NUMBER: 08/049,254
; EARLIER FILING DATE: 1993-05-15
; EARLIER APPLICATION NUMBER: 08/410,602
; EARLIER FILING DATE: 1995-04-24
; EARLIER APPLICATION NUMBER: 08/472,934
; EARLIER FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 14
; LENGTH: 1597
; TYPE: PRT
; ORGANISM: Mus musculus
US-08-628-829-14

Query Match 6.9%; Score 303.5; DB 4; Length 1597;
Best Local Similarity 22.2%; Pred. No. 2.5e-15;
Matches 113; Conservative 99; Mismatches 165; Indels 131; Gaps 23;
Qy 296 PGSKPSGNDSCBLNLSKVNSHFKPELVDSKSELIITDSTLKNKTESLLAKLE-- 353
Db 1194 PG--PGGGDSVPAPKPVNT-----APDTRGSSVPENDRL-----ASTAAELQFR 1234
Qy 354 -----ETKEYQPEVPESEN-----OKQWQAKRKSCINONPAASNNHWOIPELARKVN 401
Db 1235 SLRSHSPTERDEPAYPRSDSSSTRSWELR---TLISQTKDSASKQGPFAIQSVR 1291
Qy 402 TEQKHTTFEQPVFSVKOSPPISTSKWFDPKSICKTPSSNTLDDYMSCFTPVVKNDFPP 461
Db 1292 -----LFEERRRYEMRKNIIIG-----QVCDTPKS--YDVMVHVLGRKVT----- 1329
Qy 462 ACQLSTPYGPACFQOQHQILATPLQLVLASSANECISVKGRISYILKQISGGSS 521
Db 1330 -----FKWQRG-----NKIGEGQYG 1344
Qy 522 KVFOVLN-EKKQIYAIKVNLEADNQTLDPSYRNEIAYLNKLOQHSDDKIIRLYDYBITDQ 580
Db 1345 KVTTCISVDTGELMAKKEIRFPNDHKTIKETADELKIFEGIKH--PNLVRYFGVELHRE 1402
Qy 581 YIYVME-C--GNIDLSNLKXKSIDPWERKSYWKNWLEAVHTIHCHGIVHSDLKPAN- 636
Db 1403 EMYIFMYCEGTLLEVSRLGLQEHV-----IRLYTKQITVAINVLEHGHVIRHDIKGANI 1458
Qy 637 FLIVDGMKLIIDFGIANQMOPDTSVVKD--SQVGTVMYMPPEAIKDMSSRENGSKSK 694
Db 1459 FLTSSGLIKLGDGFCVSKLNNNAQTMEGVNSTLGTAAAYMAPEVI-----TRAKGEHGR 1513
Qy 695 ISPKSDVMSLGCILYNYTYGKTPFQOIINOISKLHAIIDPNHEIEPDIPEK---DLQDV 751
Db 1514 ---AADISLGCVVIVMTGKRPWHEYEHNFQIMYK-VGMGHK---PPPERLSPEGKAF 1566
Qy 752 LKCCLRDPKQKORISIPPELLAHYVQIQ 779

```

Db 1567 LSHCLESDPKIRWTASQLLDHAFVKVCT 1594

RESULT 14
US-09-729-995-2
; Sequence 2, Application US/09729995
; Patent No. 6426206
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO00904
; CURRENT APPLICATION NUMBER: US/09/729,995
; CURRENT FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Human
US-09-729-995-2

Query Match 6.8%; Score 298.5; DB 4; Length 505;
Best Local Similarity 28.8%; Pred. No. 1.2e-15;
Matches 92; Conservative 58; Mismatches 109; Indels 61; Gaps 14;

QY 493 LASSANECISVKGRIYSILKQIGSGSSKVFQVLNEKK-QIYAIKYVNLEADNQ----547
DB 114 VAISDAEDCVQLNQ--YKLQSEIGKGYGVVRLAYNESEDRHYAMKVLKLLKQYGF 171
QY 548 -----TLDYSRNEIAYLNKLOO-HSDKIIRLYDYEITDQYIYWM 586
DB 172 RRPPIRGSAQAQGGPAKQLPLFLERVYQEIAILKKLDHVNKKLIEVLDDPAEDN-LYLVF 230
QY 587 ECGNIDLNSWLK-----KKSIDPWERKSYKMNLEAVHTIHOHGIVHSDLKPNFL 638
DB 231 DL-----LRKGPVMEVPCDFESEQARLYLRDVLGLEVYHCKIVHRDIKPSNLL 282
QY 639 I-VDGMLKLIIDFGIANQMOPDTTSVVKDSQVGTVYNYMPPPAIKDMSSSRENGSKSKISP 697
DB 283 LGDDGHVKIADFGVSNQFEGNDAQL--SSTAGTAPAFWAPPAISD-SGQSFSGKA-----333
QY 698 KSDVWSLGCILYYMTYKTFPQOIINQISKLHAIIDPNHEIEPDIPE--KDLQDVLKCC 755
DB 334 -LDVWATGVTLYCFVYKCPFP--IDDFILALHRKI-KNEPVVPPEPEISEELKDLILKM 389
QY 756 LKRDPKQKISIPPELLAHPYV 775
DB 390 LDKNPETRIGVDPDKLHPWV 409

Search completed: September 29, 2004, 16:17:26
Job time : 39 secs

```

```

; ORGANISM: Homo sapiens
US-10-135-689-2

Query Match 6.8%; Score 298.5; DB 4; Length 505;
Best Local Similarity 28.8%; Pred. No. 1.2e-15;
Matches 92; Conservative 58; Mismatches 109; Indels 61; Gaps 14;

QY 493 LASSANECISVKGRIYSILKQIGSGSSKVFQVLNEKK-QIYAIKYVNLEADNQ----547
DB 114 VAISDAEDCVQLNQ--YKLQSEIGKGYGVVRLAYNESEDRHYAMKVLKLLKQYGF 171
QY 548 -----TLDYSRNEIAYLNKLOO-HSDKIIRLYDYEITDQYIYWM 586
DB 172 RRPPIRGSAQAQGGPAKQLPLFLERVYQEIAILKKLDHVNKKLIEVLDDPAEDN-LYLVF 230
QY 587 ECGNIDLNSWLK-----KKSIDPWERKSYKMNLEAVHTIHOHGIVHSDLKPNFL 638
DB 231 DL-----LRKGPVMEVPCDFESEQARLYLRDVLGLEVYHCKIVHRDIKPSNLL 282
QY 639 I-VDGMLKLIIDFGIANQMOPDTTSVVKDSQVGTVYNYMPPPAIKDMSSSRENGSKSKISP 697
DB 283 LGDDGHVKIADFGVSNQFEGNDAQL--SSTAGTAPAFWAPPAISD-SGQSFSGKA-----333
QY 698 KSDVWSLGCILYYMTYKTFPQOIINQISKLHAIIDPNHEIEPDIPE--KDLQDVLKCC 755
DB 334 -LDVWATGVTLYCFVYKCPFP--IDDFILALHRKI-KNEPVVPPEPEISEELKDLILKM 389
QY 756 LKRDPKQKISIPPELLAHPYV 775
DB 390 LDKNPETRIGVDPDKLHPWV 409

Search completed: September 29, 2004, 16:17:26
Job time : 39 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 16:11:35; Search time 46 Seconds
(without alignments)

1758.632 Million cell updates/sec

Title: US-10-081-119-14

Perfect score: 4376
Sequence: 1 MNKVRDKNKFKNEDLTDEL.....GGESHSSSSKTFEKKRKK 841

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PIR 78.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4376	100.0	857	2 A42861	protein kinase TTK
2	3182.5	72.7	830	2 P44439	protein kinase (EC
3	3165.5	72.3	856	2 A44439	protein kinase (EC
4	673.5	15.4	678	2 T43539	spindle checkpoint
5	643	14.7	720	2 A96807	hypothetical prote
6	586.5	13.4	764	1 S67561	protein kinase RPK
7	434.5	9.9	536	2 P90132	hypothetical prote
8	326	7.4	460	2 S58882	protein kinase Cds
9	324.5	7.4	576	2 T41587	probable carbon ca
10	321	7.3	587	2 JC5639	Ca2+/calmodulin-de
11	318.5	7.3	1607	2 T03022	MAP kinase kinase
12	317	7.2	1016	2 T25433	hypothetical prote
13	316.5	7.2	417	2 JE0191	calcium/calmodulin
14	309	7.1	1004	2 H88562	protein C07A9.3 [i
15	307.5	7.0	1044	2 S40704	hypothetical prote
16	304.5	7.0	1401	2 T39225	MAP kinase kinase
17	300	6.9	505	2 A57156	Ca2+/calmodulin-de
18	299	6.8	1478	2 S20117	protein kinase BCK
19	297	6.8	746	2 S62365	SNF1-related prote
20	296.5	6.8	357	2 T37317	probable Ca2+/calm
21	296.5	6.8	1097	2 P96538	hypothetical prote
22	295	6.7	1080	2 S48944	hypothetical prote
23	294.5	6.7	515	2 A25048	regulatory protein
24	294.5	6.7	631	2 A57286	probable serine/th
25	294.5	6.7	1081	2 S51899	probable protein k
26	294.5	6.7	1142	2 S50632	protein kinase PAK
27	293.5	6.7	1338	2 T30565	MAP kinase kinase
28	291	6.6	465	2 B55748	protein kinase (EC
29	291	6.6	925	2 A55748	protein kinase (EC

ALIGNMENTS

RESULT 1

A42861
protein kinase TTK (EC 2.7.1.-) - human
N:Alternate names: phosphotyrosine picked threonine kinase (PYT)
C:Species: Homo sapiens (man)
C:Date: 10-Jun-1993 #sequence revision 18-Nov-1994 #text_change 03-Nov-2000
R:Accession: A42861; S27971; I38144
R:Miller, G.B.; Schmandt, R.; McGill, M.; Amendola, A.; Hill, M.; Jacobs, K.; May, C.; R
J. Biol. Chem. 267, 16000-16006, 1992
A:Title: Expression of TTK, a novel human protein kinase, is associated with cell proli
A:Reference number: A42861; MUID:92348472; PMID:1639825
A:Accession: A42861
A:Molecule type: mRNA
A:Residues: 1-857 <MIL>
A:Cross-references: EMBL:M86699; NID:G340010
A>Note: sequence extracted from NCBI backbone (NCBI:109875, NCBI:109876)
A:Note: it is uncertain whether Met-1 or Met-17 is the initiator
R:Miller, G.B.; Schmandt, R.; McGill, M.; Amendola, A.; Hill, M.; Jacobs, K.; May, C.; R
submitted to the EMBL Data Library, February 1992
A:Description: Expression of TTK, a novel human protein kinase.
A:Reference number: S27971
A:Accession: S27971
A:Molecule type: mRNA
A:Residues: 17-857 <MIL2>
A:Cross-references: EMBL:M86699; NID:G340010; PIDN:AAA61239.1; PID:G340011
R:Lindberg, R.A.; Fischer, W.H.; Hunter, T.
Oncogene 8, 351-359, 1993
A:Title: Characterization of a human protein threonine kinase isolated by screening an
A:Reference number: I38144; MUID:93149596; PMID:7678926
A:Accession: I38144
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 'MCMR', 504-767, 'V', 769-802, 'GI', <LIN>
A:Cross-references: EMBL:X70500; NID:G312815; PIDN:CAA49912.1; PID:G312816
C:Genetics:
A:Gene: GDB:TTK
A:Map position: 7p12-7cen
A:Superfamily: protein kinase homology
C:Keywords: phosphotransferase
F:523-791/Domain: protein kinase homology <KIN>

Query Match 100.0%; Score 4376; DB 2; Length 857;
Best Local Similarity 100.0%; Pred. No. 2e-170;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKVRDKNKFKNEDLTDELSTLNKISADTNSGTNQMANNPEWLSLLKLEKNS 60
DB 17 MNKVRDKNKFKNEDLTDELSTLNKISADTNSGTNQMANNPEWLSLLKLEKNS 76
QY 61 VPLSDALNKLGRYSQAIEALPPDKYGNESFARIQVRFAELKAIQEPDDARYFQMAR 120

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999
C/Accession: A44339; S27865
R/Douville, E.M.; Afar, D.E.; Howell, B.W.; Letwin, K.; Tannock, L.; Ben-David, Y.; Pawa
Mol. Cell. Biol. 12, 2681-2689, 1992
A/Title: Multiple cDNAs encoding the esk kinase predict transmembrane and intracellular
A/Reference number: A44339; MUID:92269841; PMID:1375325
A/Accession: A44339
A/Molecule type: mRNA
A/Residues: 1-856 <DOU1>
A/Cross-references: EMBL:M86377; NID:g193109; PIDN:AAA37578.1; PID:g193110
A/Experimental source: p19 embryonic carcinoma cell line
A/Note: sequence extracted from NCBI backbone (NCBIN:103055, NCBIPI:103061)
C/Superfamily: protein kinase homology
C/Keywords: ATP; phosphotransferase; serine/threonine/tyrosine-specific protein kinase
F:522-790/Domain: protein kinase homology <KIN>

Query Match 72.3%; Score 3165.5; DB 2; Length 856;
Best Local Similarity 71.8%; Pred. No. 2.2e-121;
Matches 622; Conservative 86; Mismatches 107; Indels 51; Gaps 8;

QY 1 MNKVRDINKFKNEDELTDLSLNKISADTTDSGTGVTNQIMMANNPDMLSLLKLEKNS 60
DB 17 MSKMRDIFNKI-NEDCTDELSLKICADHTE--TVNQIMKVGNTPENWLNFLFKLEKNS 72
QY 61 VPLSDALLNLKLGYSQAIETALPPDKYQNESFARIQVRFPAELKAIQEPDDARDYFQMAR 120
DB 73 SPLNDLLNLKLGYSQAIETVLPDKYQNESFARIQVRLAELKAIQEPDDARDYFQMAR 132
QY 121 ANCKKFAFVHISPAQFELSGNVKYSKQILQKAVERGAVPLEMLEIALNLNLQKKLLS 180
DB 133 ENCKKFAFVHISPAQFELSGNVKYSKQILQKAVETGAVPLQMLETAMRNLHQKKQLP 192
QY 181 EEEKKNLSASTVLTAQESFSGSLGHLQNRNNSCDRGOTTKARELYGENNPPQDAEIGYR 240
DB 193 EEDKKSVSASTVLSAQBPFFSSLGWQNRNNSICESRGQAGARVLYGENLPQDAEVRHQ 252
QY 241 NSLRQTNKTKQSCFGRVPVNLNSPCDVKTTDSV-----VPCSMKR 283
DB 253 NPFQTHAAKRSCFGRVPVNLNSPFYVKTDSAVTQLTTRLALSSVPYPVYTCILHL 312
QY 284 Q-----TSRSECDRLVVGSPKSGNDSCELANLKSQNSHFEKPLVDSKESSELLIT 335
DB 313 QLLALAGLAKSGGPDRAILGSPRGRSDSYELGSLPIQTIYLDKSLVSKESSEL-MS 371
QY 336 DSITLKNKTESSLLAKLEETKEYQEPVPEPNQKQAKRSECIQNQNPAASSHWOIPE 395
DB 372 DLIALKSKTDSS-LTKLEETK---PEIAERRPPQWQSTRKPCVCFQNPAAFLRHPVD 426
QY 396 LARKVNTQKHITTEQPVFVSQSPPISTSKWDPKSI CKTPSNTLLDDYMSCFRPVW 455
DB 427 VTPKAD-----KESPPISVPKWLDPKSACTETPSSSLDDYMKCFKTPVW 470
QY 456 KNDFFPACQLSTPYGQACFOQOQHIIATPLQNLVLASSANECISVKGRISYILKQI 515
DB 471 KNDFFPACPSSTPYSQLARLQOQOQGLSTPLQSIQISGSSSINECISVNGRIYSILKQI 530
QY 516 GSGSSKVFQVLNKKQIYAIKYVNLEADNQTLDYSYRNEIAYLNKLOQHSKDIIRLYDY 575
DB 531 GSGSSKVFQVLNKKQINAIKYVNLEDADSGTIESYRNEIAYLNKLOQHSKDIIRLYDY 590
QY 576 EITQYIYVMWECGNIDLSNMLKKKSDPNERKSYWKWMLAEVHTIHQHGIVHSDLKPA 635
DB 591 EITQYIYVMWECGNIDLSNMLKKKSDINPWRKSYWKWMLAEVHTIHQHGIVHSDLKPA 650
QY 636 NFLIVDGLMLKIDFGIANQMOFDTTSVVYKDSQVGTWYTMPPEAIKDMSSSSRENGSKSKI 695
DB 651 NFVIVDGLMLKIDFGIANQMOFDTTSIVKDSQVGTWYMAPEAIRDMSSSSRENSKIRTKV 710
QY 696 SPKSDVNSLGCILYMYTGTPTFQIINOISKLHAIIDPNEHEIFEPDIPKDLQDVLKCC 755
DB 711 SPKSDVNSLGCILYMYTGTPTFQIINOISKLHAIINPAHEIPEPEISEKDLRVLKCC 770
QY 756 LKRPDKORISIPELLAAHPYVQIQTHPVNMQAKGTTTEEMKYVLGQVLGNSPNSILKAAT 815

Query Match	15.4%;	Score	673.5;	DB 2;	Length	678;
Best Local Similarity	30.9%;	Pred. No.	1.6e-20;			
Matches	195;	Conservative	102;	Mismatches	210;	Gaps 233;
Qy	233	QDAEIGYRSLRQTNTKTKQSCPFGRVFPVNLNLSPPDCDKVTDDSVVPCFMKQTSRSECRD	292			
Db	32	QDPELYFKN---DTFSSKSS-----HSDGTVTGDTLRFQSSGATALE	70			
Qy	293	LWV--PSGK-----PSGNDSCELR--NLKSVQNSHFKPELVSDSKSELIITDITLKNK	343			
Db	71	RLVSHPTKFNFDLOGNGQNSALKEVNTPAYQSMHFEHLITPLPS-----TNASH	121			
Qy	344	TESSLLAKLBETKEYQOPE-VPESNQK-----QWQAKRKSECIINONPAASNHHQIPELA	397			
Db	122	SEVSLSGAVNDLNSSEHDLPLPKVNWTKPGSLISIRRRRIGRIGLGP-----PKRA	172			
Qy	398	R-----KNTQEKHTTFQOPVESVKQSPPISTS---KWFDPKSICKTSPSSNTLDDYM	447			
Db	173	EYTLTDPKSTSDTKNSTEADEDIEMKGREVSPAGNSVAATLTKPLQLHNTPLQTSQEHPK	232			
Qy	448	SCFRTPVVKNDFPPACOL-----STPYGQP-ACFQQQQ-----HQI-----	482			
Db	233	PSFHPQSOFESFSRVOFDHDEVERRASELHSRPVTVPQEPORSASQPYESHALSFKVAPL	292			
Qy	483	----LATPLQNLQVLASSANECISVKGRIVYILKQIGGGSGSKVFOVLN-EKKQIYAIK	537			

```

Db 293 FDNQATPIPKRQ-----QDVVTANLQIKLGVGVKGGSSMVYRIFSPDNRSLYALK 345
QY 538 YVNLEADNQTLDYRNEIAYLNKLOHSDKIIRLYDEYITDQ--YIYMWECGNTIDLNS 595
Db 346 EVNFINADDTIQYKNEIALRLK-SGNDRIKLYAAEVNDITLQNLNWMWECGETDLAN 404
QY 596 WLKK--KXSIDPWRKSYKMKMLBAVHTIHQGHVHSDLKRPANFLIVDGMKMLIDFGIAN 653
Db 405 LLKMKMKPINLIRMYWECMLEAVQVHDQNIHSDLKRPANFLIVDGMKMLIDFGIAK 464
QY 654 QMQPDTTSVVKDSQVGTVMYPPAIKDMSSSRENGSKSKISPKSDVNSLGCILLYMTY 713
Db 465 AIGNDTNIHRDISHIGITINYPALTDMAHTNSGVKLVKLGKRPDSVNSLGCILLYQMY 524
QY 714 GKTPFOQIINQISKLHAIIDPNHEIEFPD-----IPEK-----DLQDVLKCC 755
Db 525 GRAPFAH-LKMIQALAIAPNEQYHIFPEVALPANAVQEXGSLPGVTYVGPULMDWVKRC 583
QY 756 LKRPDKORISIPELLAPHYVQIQTHFVNQAK 787
Db 584 LERDQRKRLTIPELLVHPFLNPLPSYLTPLAK 615

RESULT 5
A96807
hypothetical protein T32E8.5 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A96807
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A96807
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-720 <STO>
A:Cross-references: GB:AE005173; NID:g6437532; PIDN:AAF08564.1; GSPDB:GN00141
A:Gene: T32E8.5
A:Map position: 1

Query Match 14.4%; Score 643; DB 2; Length 720;
Best Local Similarity 28.1%; Pred. No. 2.9e-19;
Matches 205; Conservative 99; Mismatches 242; Indels 184; Gaps 24;

QY 107 QEPDARDYQMARAN-----CKKAFVHISFAQFELSGVNSKQVNSKQLQKAVRGAVP 160
Db 30 QRSDVVYTLSSQAANTLTQDTNNLAITSVAGESASITQPTVSEHFNDSRQMDPGKSA 89
QY 161 LEMLEIALRNILNQKQLLSEEEKNLASASTVLTAE-----SFGSLGLHQLNRNNSC-D 214
Db 90 VTSLE-----SNLDVQRK-----SQSLGTSQDMEDATNQAEASHL-----DACIG 131
QY 215 SRQTTKARELYGENPPQDAETGYRNSLRQTNK-TKQSCPGRVPVNNLNSPCDCKVTD 273
Db 132 SRKQ-----NLPVSDSEVLSKSEYKDSLSLAKIQQLGEFP-NFLNQP----- 173
QY 274 DSVVPCFMKQTSRSCRDILWPGSKPSGNDSCELRNLSQVNSHFKPELVSDKSELI 333
Db 174 -----RTRC-----SAVGSNAITLLI 190
QY 334 ITDSITLKNKTESSLLAKLEETKEYQEPPEVPSNQKQWAKRSECTINQNPAASSHQWI 393
Db 191 HSSSAPMLNAT-----THVRSYVEAD-----SNANPHAVQSQGNL 226

```

```

QY 394 PELARKVNTQKHTFEQVFSVSKQSPPISTKSWFDPKSICTKPSSTNTLDYMSCRTP 453
Db 227 PSCC-----PSSKVSNIHPNKDATASEMPASTN---DPEVVKETDTSKQOQITGLEAP 279
QY 454 VVNDFFPACQLSTPYQP-----ACFOOQHOILATPLQNL---QVLIASSANEC----- 501
Db 280 VG-----SSYIGSDGQANARLPEELHTSVSSOPQKSDRHEKVASSKGPAPAKRN 329
QY 502 -----ISVKGRIYSILKQIGSGSSKVFQVNLNEKKQIYAIKYVNLBEADNQTLDSYRNE 555
Db 330 YDPDLFFKNGKLYORLKIIGSGSSSEVHKVVISSDCTIYALKKIKLKGKRDVATAYGCOE 389
QY 556 TAYLNKLOHSDKIRLYDYBITDQ-----YIYMWECGNTIDL 593
Db 390 IGYLKLKGTN-IIQLIDYEVTDKTLQEVNLGTMNSKDGKVEDGFIYVWLEYGIDL 448
QY 594 NSWLKKK-----KSTDPWRKSYKMKMLBAVHTIHQGHVHSDLKRPANFLIVDGMK 645
Db 449 AHMUSQKWRIEGSDRTIDENLRFYWOQIILQAVNTIHERIVHSDLKRPANFLIVRGFLK 508
QY 646 LIDFGIANQMPDTSVVKDSQVGTVMYPPAIKDMSSSRENGSKSKISPKSDVNSL 705
Db 509 LIDFGIAKAINSDTTNIQRDSQVGTLSYMSPEAFW-CNESDENGNT-IKCGRPSDINS 566
QY 706 CILYMTYKTPFQOINQISKLHAIIDPNHEIEFPDIPKQLODVLKCLKRPDKORIS 765
Db 567 CILYQMYVGRTPPFADYKTFWAKFKVITDPNHEITYNQLSNPWLIDLMKKCLAWDRNQWR 626
QY 766 IPELLAHPIYV 775
Db 627 IPELLQHPFL 636

RESULT 6
S67561
protein kinase RPK1 (EC 2.7.1.1) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein D2785; protein YDL028c; spindle pole body duplication protein
C:Species: Saccharomyces cerevisiae
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
R:Paulin, L.; Saren, A.M.; Laamanen, P.
A:Accession: S67561; S46425; S55256
Submitted to the Protein Sequence Database, July 1996
A:Reference number: S67560
A:Accession: S67561
A:Molecule type: DNA
A:Residues: 1-764 <PAU>
A:Cross-references: EMBL:D274076; NID:g1431003; PIDN:CAA98587.1; PID:g1431004; GSPDB:GN00
A:Experimental source: strain S288C
R:Poch, O.; Schwob, E.; de Fraipont, F.; Camasses, A.; Bordonne, R.; Martin, R.P.
Mol. Gen. Genet. 243, 641-653, 1994
A:Title: RPK1, an essential yeast protein kinase involved in the regulation of the onset
A:Reference number: S46425; MUID:94301294; PMID:8028550
A:Accession: S46425
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-145; 'S', 147-210; 'TKR', 214-764 <POC>
A:Cross-references: GB:L08909; NID:g1197057; PIDN:AAA88731.1; PID:g1197058
R:Laube, E.; Stoelcker, B.; Luca, F.C.; Weiss, E.; Schutz, A.R.; Winey, M.
EMBO J. 14, 1855-1863, 1995
A:Title: Yeast spindle pole body duplication gene MPS1 encodes an essential dual specific
A:Reference number: S55256; MUID:95255223; PMID:7737118
A:Accession: S55256
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 446-491; 'IERP', 497-715 <LAU>
C:Genetics:
A:Gene: SGD:MPS1; RPK1; MIPS:YDL028c
A:Cross-references: SGD:S0002186; MIPS:YDL028c
A:Map position: 4L
C:Superfamily: yeast protein kinase RPK1; protein kinase homology
C:Keywords: phosphotransferase
F:438-716/Domain: protein kinase homology <KIN>

```


A;Description: EC 2.7.1.-; protein kinase Cds1 [validated, MUID:98119835]; is required b
 A;Note: Cds1 is phosphorylated and activated by S-phase arrest and activated by DNA dama
 C;Superfamily: protein kinase Cds1; kinase interaction domain homology; protein kinase h
 C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
 F;62-133/Domain: kinase interaction domain homology <KIH>
 F;165-433/Domain: protein kinase homology <KIN>
 F;173-181/Region: protein kinase ATP-binding motif

Query Match 7.4%; Score 326; DB 2; Length 460;
 Best Local Similarity 24.4%; Pred. No. 1.2e-06;
 Matches 122; Conservative 95; Mismatches 196; Indels 88; Gaps 22;

Qy 322 PL-VSDSKSSELIITDITKNTKTESLLAKLSETEYQEPVPSNQWQAKRSEK- 379

Db 15 PLVYSQNIKQVNNENFVK- - - - -LVMTMLDGKTEVIPLTVDVHGFWRFGHKSCE 69

Qy 380 -INQNPAASNHHQIPELARKVATEOKHTTFEQPVFSVSKQSPPISTSKWFDPKSICKTP 438

Db 70 VLNGPRVSNFHEIYQGRHNDSENVVP- - - - -LHDHS- - - - -SNGTFLNFERLAK- 118

Qy 439 SSNTL- - - - -DYMSCFPTPVVKNDFPPACQLSTPYGQACFOQQQHQIILATPLQNLQVLA 494

Db 119 NSRTILSGDEIRIGLGP- - - - -KDBISPLCQVPVKHSR- - - - - 153

Qy 495 SSSANECSVKGRIYSILKQIGSGSSKVPQVLNKKQI- - - - -YAIKYVN- - - - -LEEA 544

Db 154 -DSQXNWKSGNSHYEIRITLGG- - - - -TPAVKLAVEVNSGKWAYIKIRKLLTSS 208

Qy 545 DNOTLDSYRNEIAYNLKLOHSHKIRLYDEITDQIYVMVE- - - - -CGNIDLSNLKKKSI 603

Db 209 EKRATEMFORIEDILKSL- - - - -HHPGVQVCHCEICNDDELFTWVEYVEGGDLMDFLIANGSI 266

Qy 604 DPWERKSYWKNMLEAVTHIQHGVHSDLPANFLIV-DGMLKLIIDFGIANQVQPTTSV 562

Db 267 DEQDCKPLQLQLETLHLHKGQVTHRDIPENILITNDPLKISDFGLAKVIH- - - - -GTGT 324

Qy 663 VKSQVGTVMVMPPEALKMSSSRENGKSKSIKSPKSDVMSLGCILYIYMTYGTTP- - - - - 718

Db 325 FLETFCTGMYLAPEVLKSNVNLDCG- - - - -YDDKVDIWSLGCVLVYVMTASIPASSS 379

Qy 719 -QQIINQISKLHAIIDPNHIEPDPPEKDLQVLKCCLRKDPKQISIPPELLAHP-YVQ 776

Db 380 QAKCIELISGAYPIEPLLENE- - - - -ISEGI-DLINRMLEINPEKRISSEALQHPWFYT 435

Qy 777 IQTH- - - - -PVNQMAKGTTEEM 793

Db 436 VSTHEHTPPSSSEHEATEQL 456

RESULT 9
 T41587
 probable carbon catabolite derepressing protein kinase - fission yeast (Schizosaccharomy
 C;Species: Schizosaccharomyces pombe
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
 C;Accession: T41587
 R;Murphy, L.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, September 1998
 A;Reference number: Z21968
 A;Accession: T41587
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-576 <MUR>
 A;Cross-references: EMBL:AL031543; PIDN:CAA20833.1; GSPDB:GN00068; SPDB:SPCC74.03c
 A;Experimental source: strain 972h-; cosmid c74
 C;Genetics:
 A;Gene: SPDB:SPCC74.03c
 A;Map position: 3
 A;Introns: 50/1; 127/3; 245/1
 C;Superfamily: AMP-activated protein kinase; protein kinase homology

Query Match 7.4%; Score 324.5; DB 2; Length 576;

Best Local Similarity 29.3%; Pred. No. 1.8e-06;

Matches 102; Conservative 67; Mismatches 132; Indels 47; Gaps 15;

Qy 509 YSILKQIGSGSSKVFQVLNEK-KQIYAIKYVN- - - - -LEEADNQTLDYSRNEIAYNLKLOQ 564

Db 34 YIIRETLGEGSGFKVLKATHYKTKQKVALKFIISQLLKKSDMHM- - - - -RVEREISYL-KLLR 90

Qy 565 HSDKIIRLYDEITDQIYVMVECGNIDLSNLKKSIDPWERKSYWKNMLEAVTHIQ 624

Db 91 H-PHIILYDVITPTDIVMWIEYAGGELFDYIYEKRMTEDEGRFRFQQIICAIEYCHR 149

Qy 625 HGVHSDLKPANFLIVDGM-LKLIDFGIANQVQPTTSVVKDQVGTVMVMPPEALKDM 683

Db 150 HKIVHREDLKPEJLLDDNLNVKIADFGLSNIM- - - - -TDGNFLKTCGSGSPNVAPEVI- 202

Qy 684 SSRENGKSKISPKSDVMSLGCILYIYMTYGTTPQ- - - - -IINQISKLHAIIDPNHIEFP 741

Db 203 - - - - -NGKLYA- - - - -GPEVDVMSQGVLYVYVNLVGRLPDDEFIPNLFKYNVCV- - - - -YVMP 251

Qy 742 DIPEKDLQVLKCCLRKDPKQISIPPELLAHPYVQI- - - - -THPVNQMAKGTTEEMKYV- 796

Db 252 DFLSFGAQLIRRMIVADPMQRIITQIERRDPFNVNLPDYLRMEV- - - - -QSVADSRIVS 310

Qy 797 -LGQLVGLNSP- - - - -NSILKAAKTLYEHYSGSHSSSSK 832

Db 311 KLGEAMGFSEDIYVLEALRSDENNEVKEAYNLLHENQVIQEKSHLSKSK 358

RESULT 10
 JC5669

Ca2+/calmodulin-dependent protein kinase kinase (EC 2.7.1.-) beta chain - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 20-Nov-1997 #sequence_revision 20-Nov-1997 #text_change 24-Sep-1999

C;Accession: JC5669; PC4493

R;Kitani, T.; Okuno, S.; Fujisawa, H.

J. Biochem. 122, 243-250, 1997

A;Title: Molecular cloning of Ca2+/calmodulin-dependent protein kinase beta.

A;Reference number: JC5669; MUID:97420710; PMID:9276695

A;Accession: JC5669

A;Molecule type: mRNA

A;Residues: 1-587 <KIT>

A;Cross-references: GB:AB018081; NID:g3702720; PIDN:BAA33524.1; PID:d1034490; PID:g37027

A;Experimental source: cerebellar

A;Accession: PC4493

A;Molecule type: protein

A;Residues: 425-501 <KI2>

A;Experimental source: cerebellar

C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C;Keywords: ATP; phosphotransferase

F;162-445/Domain: protein kinase homology <KIN>

F;170-194/Domain: ATP-binding #status predicted <ATP>

F;480-493/Domain: calmodulin-binding #status predicted <CAB>

Query Match 7.3%; Score 321; DB 2; Length 587;

Best Local Similarity 25.5%; Pred. No. 2.6e-06;

Matches 100; Conservative 80; Mismatches 126; Indels 86; Gaps 19;

Qy 419 QSPFISTKMFDPKSIKCTPSSNTLDDYMSCFRTFVVVNDFFPACQLSTPYGQACFPQ- 477

Db 105 QGGPASSSS-LDMNGRCICPSLS- - - - -YSPA- - - - -SSPSSSPRMPRRP 143

Qy 478 - - - - -QQHQILATPLQNLQVLASSANECSVKGRIYSILKQIGSGSSKVFQVLNEK-KQIY 534

Db 144 TVSEHHVSTIGLQ- - - - -DCVQLNQ- - - - -YTLKDEIGKSGYGVVKLAYNENDNTYY 190

Qy 535 AIKYVNLEEAADNQ- - - - -LDSYRNEIAYNLKLOHSD- - - - -K 568

Db 191 AMKVLKSKLIRQAGFPFRPPRGRTRPAGGCIQPRGPIEQVYQETAILKKL-DHPNVVK 249

Qy 569 IIRLYDEITDQIYVMVECGNIDLSNLKKSIDPWERKSYWKNMLEAVTHIQHIV 628

Db 250 LVEVLD-DPNEHDHLYMVVFELVNOGVNMEVPTLPSEDAQRFYQDLKIGIEYHLQKLI 308

Qy 629 HSDLKPANFLI-VDGMLKLIIDFGIANQVQPTTSVVKDQVGTVMVMPPEALKDMSSSRE 687

Db 309 HRDIKPSNLLVGEDGHKIAIDFGVSNFEK--GSDALLSNTVGTGPAFWAPES---LSETRK 363
Qy 688 --NGKSKSKSPKSDVNSLGCILLYMTYKTPQOIIHQIINQISKLHAIIDPNHEIEFPDPIPE 745
Db 364 IFSGKA-----LDVWANGVTLVCFVFGQCPMD--ERIMCLHSKI-KSQALEFPDQPD 413
Qy 746 --KDLQVLKCKLRDPKQISIPPELLAHPVV 775
Db 414 IAEDELKLIITMLDKNPSRIVPEIKLHPV 445

RESULT 11
T03022
MAP kinase kinase - human
C:Species: Homo sapiens (man)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 05-Nov-1999
C:Accession: T03022
R:Takekawa, M.; Posas, F.; Saito, H.
submitted to the EMBL Data Library, May 1997
A:Description: A human homolog of the yeast Ssk2/Ssk22 MAP kinase kinase, MTK1, a
A:Reference number: Z14824
A:Accession: T03022
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1607 <FAK>
A:Cross-references: EMBL:AF002715; NID:G2352276; PIDN:AAB68804.1; PID:G2352277
C:Genetics:
A:Note: MTK1

Query Match 7.3%; Score 318.5; DB 2; Length 1607;
Best Local Similarity 23.0%; Pred. No. 9.6e-06;
Matches 116; Conservative 98; Mismatches 161; Indels 129; Gaps 23;

Qy 300 PSNGDSCELRNKLVQNSHPKEPLVSDKSESLIITDITLKNKTESLLAKLE-----353
Db 1206 PSQDSV--LPKISSAH-----DTRGSSVPENDRL-----ASIAELQFRSLR 1248
Qy 354 --ETKQYQPEVPE-----SNQWQAKRKSECCINQPAASNHNQIPELARKYNTQEK 405
Db 1249 HSSPTERDEPAYRGDSSGSTRSWELR---TLISQSKDTASKLGPPIEAQKSVR---1301
Qy 406 HTTEQPVESVSKSPISITKWFDPKSICTPTSSNTLDDYMCFTFPVVKNDFFPACQL 465
Db 1302 --LFEERYEMRKNIG-----QVCDTFKS--YDNVMVGLRKVT-----1339
Qy 466 STPYGPACFQQQHQHILATPLQNLQVLASSANECISVKGRYISILKQIGSGSGSKVFQ 525
Db 1340 -----FKWQRG-----NKIGEGYQKVT 1358
Qy 526 VLN-EKKQIYAIKYVNLLEADNQLDSYRNEIAYLNKLQHSKDIIRLYDYETDQYIYM 584
Db 1359 CISVDTGELWANKIEIRFPQNDHKTIKETADELKIIFEGIKH--PNLVRYFQVGLHREMYI 1416
Qy 585 VME-C--GNIDLNSWLKKKSIDPWERKSYKQNMLEAVHTIHOHGIVHSDLPKAN-FLIV 640
Db 1417 FMEYCDSEGTLEEVSRGLQEHV-----IRLYSKQITIAINLVEHGIHVHRDIKGANIFLTS 1472
Qy 641 DGMKLIDFGIANQMDPTTSVVKD--SQVGTVNYMPEAIKDMSSSRENGSKSKISPK 698
Db 1473 SGLIKLDFGCSVLKNQNAQTMPEVNSLTGTAAYVAPEVI-----TRAKGEHGR--A 1524
Qy 699 SDVNSLGCILLYMTYKTPQOIIHQIINQISKLHAIIDPNHEIEFPDPIPEK---DLQDLVKCC 755
Db 1525 ADINSLGCVIEMVTGKRPHEYEHNFQIMYK-VGMGHK---PPIPERLSPEGKDFLSHC 1580
Qy 756 LKRPDKORISIPPELLAHPVYQIOT 779
Db 1581 LESDPKRWNTASQLLDSFVKVCT 1604

RESULT 12

T25433
hypothetical protein ZK524.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T25433; T27888
R:McMurray, A.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z20033
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1016 <WIL>
A:Cross-references: EMBL:Z72517; PIDN:CAA96698.1; GSPDB:GN00019; CESP:ZK524.4
A:Experimental source: clone T28F4
R:Gardner, A.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z20435
A:Accession: T27888
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1016 <W12>
A:Cross-references: EMBL:Z73912; PIDN:CAA98148.1; GSPDB:GN00019; CESP:ZK524.4
A:Experimental source: clone ZK524
C:Genetics:
A:Gene: CESP:ZK524.4
A:Map position: 1
A:Introns: 52/3; 91/3; 151/3; 207/1; 254/2; 288/3; 422/2; 560/2; 592/1; 623/1; 660/3; 6

Query Match 7.2%; Score 317; DB 2; Length 1016;
Best Local Similarity 29.2%; Pred. No. 6.7e-06;
Matches 81; Conservative 64; Mismatches 102; Indels 30; Gaps 10;

Qy 508 IYSILKQIGSGSSKVFQVLEKKQIY-----AIKYNLEADNQTLSYRNEIAYLNKL 562
Db 26 LDYLEKTIQGGH-----FAVVKLAKHVTGEMVAVKIIDTKRMDEASTQIMKEVRCM-KL 80
Qy 563 QQHSKDIIRLYDYETDQYIYVMCEGNIIDLSW-LKKKSIDPWERKSYKQNMLEAVHT 621
Db 81 VQFAN-IVRLVEVLDTQKIFLLELGDYDLHDFIKHEKGVCESLAQYFCQIMTADY 139
Qy 622 IHQHGIVHSDLPKANFLIV--GMLKLIDFGIANQMDPTTSVVKDSSQVGTVNYMPEPAI 679
Db 140 CHQLHWVRDLKPNVWFPEKLGWVKLTDFGFSNSYEPGEQ---LNTSGSLAYSAPIL 196
Qy 680 KMSSSRENGSKSKISPKSVNSLGCILLYMTYKTPQOIIHQIINQISKLHAIIDPNHEIE 739
Db 197 -----LQDSYDAPADVMSLGVILYMLVCGRLPFQE-ANDSETLTILDCCKYSI- 244
Qy 740 FPDIEKDLQVLKCKLRDPKQISIPPELLAHPVVQ 776
Db 245 -PDVLSDCRNLIQSMVLRPEPKRASLEKIVSTSWVQ 280

RESULT 13
JE0191
calcium/calmodulin-dependent protein kinase kinase (EC 2.7.1.1) - human
N:Alternate names: CaMKK
C:Species: Homo sapiens (man)
C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 16-Dec-1998
C:Accession: JE0191
R:Hsu, L.S.; Tsou, A.P.; Chi, C.W.; Lee, C.H.; Chen, J.Y.
Biomed. Sci. 5, 141-149, 1998
A:Title: Cloning, expression and chromosomal localization of human Ca2+/calmodulin-depe
A:Reference number: JE0191
A:Molecule type: mRNA
A:Accession: JE0191
A:Residues: 1-417 <HSU>
C:Comment: This enzyme functions as a key intermediaries in Ca2+/calmodulin-driven sign
C:Genetics:
A:Gene: CaMKK
A:Map position: 12
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol
C:Keywords: phosphotransferase
F:47-330/Domain: protein kinase homology <KIN>

Query Match	7.28; Score 316.5; DB 2; Length 417;	
Best Local Similarity	25.78; Pred. No. 2.7e-06;	
Matches	90; Conservative 79; Mismatches 116; Indels 65; Gaps 16;	
Qy	461 PACLSFYGPQACFOQ--CQHQLATPLQNLQVLASSANECISGVKRIYSILKQIGS 517	
Db	11 PYSVSSPQSSPRLPRPTVESHVSTGQ-----DCVQLNQ--YTLKDEICK 57	
Qy	518 GGSKVFQVINEK-KQIYAIKYNLEADNQT-----LDSY 552	
Db	58 GSYGVWKLAYENDNTYAMKVLSKKLIRQAGPRPRPRGTRPAPGCIQPRGFIEQV 117	
Qy	553 RNEIAYLNKQQSD--KIIRLVDEYITDOYIYMMWECGNIDLNSMLKKKSIDPWERS 610	
Db	118 YQETAILKKL-DHPNVVLEVID-DNEDHLNVFELVNOGPMVEPLKPLSEDAQRF 175	
Qy	611 YKXMLEAVHTIHOHGVHSDLPANFLI--VDGMLKLIDFGIANQMPDTTSVVKDSQVG 669	
Db	176 YFQDLIKGIEYLHYQKLIHRDIXPSNLLVGEDGHIKIADFGVSNFK--GSDALLSNTVG 233	
Qy	670 TVNVMPEAIKDVSSSE--NGKSKSKISPKSDVWSLGCILYVMTYKTPFQOIQINQISK 727	
Db	234 TPAPMAPES---USETKIFSGKA-----LDVWAGVTLVCFVFGQCPFMD--BRIMC 281	
Qy	728 LHAIIDPNHIEFPDIPE--KDLQDVLKCLKRPQKQISIPPELLAHYV 775	
Db	282 LHSKI-KSQALEPDPQDIAEDLKDLITRMLDKNPESRIVVPBKLHPW 330	
RESULT 14		
H88562	protein C07A9.3 [imported] - Caenorhabditis elegans	
C:Species:	Caenorhabditis elegans	
C:Date:	10-May-2001 #sequence_revision 10-May-2001 #text_change 24-Aug-2001	
C:Accession:	H88562	
R:anonymous:	The C. elegans Sequencing Consortium.	
Science 282, 2012-2018, 1998		
A:Title:	Genome sequence of the nematode C. elegans: a platform for investigating biology	
A:Reference number:	A75000; PMID:99069613; PMID:9851916	
A:Note:	see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.ele	
A:Note:	published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and	
A:Accession:	H88562	
A:Status:	preliminary	
A:Molecule type:	DNA	
A:Residues:	1-1004 <STO>	
A:Cross-references:	GB:chr_III; PIDN:CAA82347.1; PID:g3873989; GSPDB:GN00021; CESP:C07A9	
C:Genetics:		
A:Gene:	C07A9.3	
A:Map position:	3	
C:Superfamily:	unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo	
Query Match	7.18; Score 309; DB 2; Length 1004;	
Best Local Similarity	20.18; Pred. No. 1.4e-05;	
Matches	176; Conservative 159; Mismatches 326; Indels 214; Gaps 39;	
Qy	44 NNPDWLSLLKLEKNSVPLSDALLNKLIGRYSAIEALPPDKYQNESFARIQVRF-AE 102	
Db	163 SSPTMQCCQAMSEDSIEMRD-----YNSGVHMHHPHQWQMQQQOHHQOQYNMS 213	
Qy	103 LKATQEPDDARDYFQMARANKKFAFVHISFAQFELSSQGNVKKSKQLLQKAVRGAVPLE 162	
Db	214 YNHQOQWQWYHQOQQOYQOQAQHHQWYAF-QIQQ-----QQQPPQSQSQSQSQPQ 268	
Qy	163 MLEIALNMLKQKLLSEKKNL-SASTVLTAAQSFSGLHQLNRNNSCDSRGQTIK 221	
Db	269 QSSAALQHN-----ESSNLSSAGSISDRPEHQGGTQRPAPQSTATDKTR 318	
Qy	222 ARFLYGENMPPQDAEIGYRNSLRNTKQSCFGRVPVNLNNSPCDVKTDSDVVPFCM 281	
Db	319 KRRKAG---PTED-----QATPKQRKIITEFMKVG-----EVASGNSVARCILL 359	
Qy	282 K--RQTSRSECDRLVPGSPKPSGNDSCELR-----NLKSVQNSHP----- 319	

Db	360 TEYHQNGSPKQ---PAVQNGSGNSYDSQQQOPMQHQMNSYWGCVATPSLGVNVRGT 416	
Qy	320 -----KEPLVSDKSSSELIITDSITLKNKTESSI-----LAK 351	
Db	417 PTPTQQHYSSDSNSNSQSPPGQNGSGRMVRT--IDEQTQDSSLSQANPQNADEVAK 474	
Qy	352 LEETKEYQEPVPVPSNOKQWQAKRKE-----CINQNPAASS----- 388	
Db	475 MNRIEDHRRRIEELNLSKNSLERRKNEASKETIKRLIDKNQIERKALDKTAADSPRIG 534	
Qy	389 -----NHWQIPELARK---VNTQKHTTTEQPVFVSVKQSP----- 421	
Db	535 CFKTRTGDSPRDQWVQWAPAEKDKKTEQINABERNEIASAL--LKTKPLGIGKEPK 592	
Qy	422 -PISTKWFDPKSICTKTPSSNTLDVMSCFRTPVVKND--PPACQLSTPY----- 469	
Db	593 RQAVNSQNSNGMQPSTSSNTNGD-DAIFRPEEPKEIQOYEVIELDEYIKLREHLRK 651	
Qy	470 -----GQPACFOQQOQHILATPLQNLQVLASSANECIS--VKGRIYSILKQIGSGS 520	
Db	652 BETDLSMEKERLEKEKHV-----RELKRASNESASQFNDHRLHLLKRYLMNLLKGGF 706	
Qy	521 SKVTOVLNEKKQIYA---IKYVNLEADNQTLDYSRNEIAY--LNKLOQSDDKIIRLYDY 575	
Db	707 SEVKAFAFIEENRYVACKIHHVKNQWKEEKKANYVKHAMREKDIHKSLDHC-RIVKQYDL 765	
Qy	576 EITQYIY-WWMEC--GNIDLSMLKKKSIDPWERSYWKNNMLEAVHTTIHQHG--IVHS 630	
Db	766 LTIDNHSFCTVLEVPGN-DLDFYVKQNRISSEKARSIIIMQVVSALVYLNEKSTPIHY 824	
Qy	631 DLKPANFLI-----VDGMLKLIDFGIANQMPDT-----TSVVKDSQ-VGTVNYMPPKAID 681	
Db	825 DLKPANILLESNGTSGAIKITDFGLSKIMGESDDHDLGLIELTSQFAGTYWYLPPTFI- 883	
Qy	682 MSSRENGKSKSPKSDVWSLGCILYVMTYGTPTPQIINO--ISKHAIIDPNHEIE 739	
Db	884 -----VPPKLTCKVDVNSIGVIFQYCIYKPKFGNDLTQOKILEYNTIINA-REVS 934	
Qy	740 FPDPIPE--KDLQDVLKCLKRPQKQISIPPELLAH 772	
Db	935 FPSKPQVSSAAQDFIRRCLOVYRKEDRADVPFLAKH 969	
RESULT 15		
S40704	hypotheical protein C07A9.3 - Caenorhabditis elegans	
C:Species:	Caenorhabditis elegans	
C:Date:	06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 28-May-1999	
C:Accession:	S40704	
R:Smith, M.	Submitted to the EMBL Data Library, December 1993	
A:Reference number:	S40701	
A:Accession:	S40704	
A:Status:	preliminary	
A:Molecule type:	DNA	
A:Residues:	1-1044 <SMI>	
A:Cross-references:	EMBL:Z25094; NID:g436440; PID:g436444	
C:Genetics:		
A:Introns:	86/3; 158/3; 265/3; 340/2; 500/3; 980/3; 1017/3	
C:Superfamily:	unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase	
C:Keywords:	ATP; serine/threonine-specific protein kinase	
F:733-1012/Domain:	protein kinase homology <KIN>	
F:741-749/Region:	protein kinase ATP-binding motif	
Query Match	7.08; Score 307.5; DB 2; Length 1044;	
Best Local Similarity	20.08; Pred. No. 1.7e-05;	
Matches	176; Conservative 159; Mismatches 326; Indels 217; Gaps 39;	
Qy	44 NNPDWLSLLKLEKNSVPLSDALLNKLIGRYSAIEALPPDKYQNESFARIQVRF-AE 102	
Db	200 SSPTMQCCQAMSEDSIEMRD-----YNSGVHMHHPHQWQMQQQOHHQOQYNMS 250	
Qy	103 LKATQEPDDARDYFQMARANKKFAFVHISFAQFELSSQGNVKKSKQLLQKAVRGAVPLE 162	

This Page Blank (uspto)

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.
 OM protein - protein search, using sw model
 Run on: September 29, 2004, 16:13:55 ; Search time 138 Seconds
 (without alignments)
 1922.832 Million cell updates/sec

Title: US-10-081-119-14
 Perfect score: 4376
 Sequence: 1 MNKVRDKNFKNEDLTDEL.....GGESHNSSSKTFEKKRGKK 841

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea:*
 2: sp_bacteria:*
 3: sp_fungi:*
 4: sp_human:*
 5: sp_invertebrate:*
 6: sp_mammal:*
 7: sp_mmc:*
 8: sp_organelle:*
 9: sp_phase:*
 10: sp_plant:*
 11: sp_rodent:*
 12: sp_virus:*
 13: sp_vertebrate:*
 14: sp_unclassified:*
 15: sp_rvirus:*
 16: sp_bacteriap:*
 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4373	99.9	857	4 Q9BW51	Q9BW51 homo sapien
2	3092.5	70.7	815	11 Q8BY97	Q8BY97 mus musculus
3	2157.5	49.3	882	13 Q98T92	Q98T92 xenopus lae
4	2077.5	47.5	585	11 Q8BX21	Q8BX21 mus musculus
5	1648.5	37.7	983	13 Q8AVG3	Q8AVG3 brachydanio
6	1646	37.6	982	13 Q7T2A4	Q7T2A4 brachydanio
7	1564	35.7	305	4 Q15272	Q15272 homo sapien
8	647.5	14.8	777	10 Q84YX4	Q84YX4 arabidopsis
9	643	14.7	720	10 Q9CA22	Q9CA22 arabidopsis
10	637.5	14.6	742	10 Q8LN55	Q8LN55 oryza sativ
11	637	14.6	630	5 Q9VEH1	Q9VEH1 drosophila
12	628	14.4	519	10 P93833	P93833 arabidopsis
13	501.5	11.5	586	5 Q8SRH4	Q8SRH4 encephalito
14	490	11.2	297	10 Q8UA38	Q8UA38 arabidopsis
15	434.5	9.9	536	10 Q9S830	Q9S830 guillardia
16	339	7.7	755	13 Q90ZY6	Q90ZY6 brachydanio

17 336 7.7 1090 5 Q9GRT3
 18 336 7.7 1090 5 Q9N9J2
 19 334 7.6 503 11 Q8BXN8
 20 334 7.6 541 11 Q8QZT7
 21 334 7.6 579 11 Q8C0G3
 22 331 7.6 588 11 Q8C078
 23 330 7.5 588 11 Q8CH42
 24 325 7.4 1429 5 Q8IHPI
 25 324 7.4 465 5 Q8T0L6
 26 324 7.4 520 5 Q9VHF6
 27 323.5 7.4 588 4 Q9VSN2
 28 322 7.4 520 5 Q86P98
 29 321 7.3 2265 5 Q8IID2
 30 320 7.3 1167 5 Q8T2I8
 31 318 7.3 1167 5 Q8T2I8
 32 317 7.2 557 4 Q84883
 33 317 7.2 1016 5 Q22855
 34 316.5 7.2 522 11 Q8KIX6
 35 316.5 7.2 522 11 Q8KIX6
 36 316.5 7.2 523 13 Q8AVR1
 37 316.5 7.2 533 4 Q96RP2
 38 316.5 7.2 533 4 Q96RR3
 39 316.5 7.2 541 4 Q8WY04
 40 316.5 7.2 541 4 Q8IUG3
 41 316.5 7.2 588 4 Q96RP1
 42 316.5 7.2 588 4 Q96RR4
 43 315 7.2 1497 5 Q95YH6
 44 315 7.2 1571 5 Q95YH7
 45 315 7.2 1571 5 Q8MSQ4

ALIGNMENTS

RESULT 1

Q9BW51 PRELIMINARY; PRT; 857 AA.
 AC Q9BW51; Q9NTMO;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE TTK protein kinase (PPT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE:Eye, and Testis;
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC000633; AAH00633.1; -
 DR EMBL; BC032858; AAH32858.1; -
 DR HSSP; Q00534; 1BI8.
 DR GO; GO:0005524; P:ATP binding; IEA.
 DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_kinase.
 DR InterPro; IPR008941; TPR-like.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Kinase.
 SQ SEQUENCE 857 AA; 97072 MW; 51F40A3CD1677AC5 CRC64;

Query Match 99.9%; Score 4373; DB 4; Length 857;
 Best Local Similarity 99.9%; Pred. No. 1.7e-270;
 Matches 840; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	1	MNKVRDIKNFKNEDELTDLSLNKI	SADTTDNGSTVQIMMANNPEDWLSLLKLEKNS	60
Db	17	MNKVRDIKNFKNEDELTDLSLNKI	SADTTDNGSTVQIMMANNPEDWLSLLKLEKNS	76
QY	61	VPLSDALLNLKIGRYSOA	TEALPPDKYQONESFARIOVPAELKATQEPDDARDYFQMAR	120
Db	77	VPLSDALLNLKIGRYSOA	TEALPPDKYQONESFARIOVPAELKATQEPDDARDYFQMAR	136
QY	121	ANCKFAFVHISFAQFELSOGN	VKSKQLLOKAVERGAVPLEMLEIALRNLNLOKQKLLS	180
Db	137	ANCKFAFVHISFAQFELSOGN	VKSKQLLOKAVERGAVPLEMLEIALRNLNLOKQKLLS	196
QY	181	EEBKXNLASVTIATAESFSG	SLGHLQNRNNSCDSRGQTTKARFLYGENNPPQDAEIGYR	240
Db	197	EEBKXNLASVTIATAESFSG	SLGHLQNRNNSCDSRGQTTKARFLYGENNPPQDAEIGYR	256
QY	241	NSLRQTNKTKQSCPFGRVP	VNLLNSPDCDVYKTDSDVVPFCMRKQTSRSECRDLVVPGSKP	300
Db	257	NSLRQTNKTKQSCPFGRVP	VNLLNSPDCDVYKTDSDVVPFCMRKQTSRSECRDLVVPGSKP	316
QY	301	SGNDSCELRLNLSQVNSHF	KEPLVSDKSESLIITDSITLKNKTESLILAKLEETKEYQE	360
Db	317	SGNDSCELRLNLSQVNSHF	KEPLVSDKSESLIITDSITLKNKTESLILAKLEETKEYQE	376
QY	361	PEVPESNQKQWAKRSECI	NQNPAASSNHQIPELARKVNTQKHITTFEQPVFSVKQS	420
Db	377	PEVPESNQKQWAKRSECI	NQNPAASSNHQIPELARKVNTQKHITTFEQPVFSVKQS	436
QY	421	PPISTSKWFPKSI	CKTPSSNTLDDYMSCFRTPVWKNDFPPACQLSTPYGQACFQOQOH	480
Db	437	PPISTSKWFPKSI	CKTPSSNTLDDYMSCFRTPVWKNDFPPACQLSTPYGQACFQOQOH	496
QY	481	QILATPLQNLQVLASSAN	ECISVKGRIVSYLKQIGSGSSKVFOVLNKKQIYAIKYN	540
Db	497	QILATPLQNLQVLASSAN	ECISVKGRIVSYLKQIGSGSSKVFOVLNKKQIYAIKYN	556
QY	541	LEADNQTLSYRNEIAYLN	KLOQSDKIIRLYDYBITDQYIYMWMECGNIDLSNLWKKK	600
Db	557	LEADNQTLSYRNEIAYLN	KLOQSDKIIRLYDYBITDQYIYMWMECGNIDLSNLWKKK	616
QY	601	KSIDPWERKSYWKNM	LEAVHTTHQGVHSDLPANFLVDGWLKIDFGIANQWQPDIT	660
Db	617	KSIDPWERKSYWKNM	LEAVHTTHQGVHSDLPANFLVDGWLKIDFGIANQWQPDIT	676
QY	661	SVVKDSQVGVNTYMP	PEAIKDMSSSRENGSKSKISPKSDVMSGLGILYMTYIGKTPFQO	720
Db	677	SVVKDSQVGVNTYMP	PEAIKDMSSSRENGSKSKISPKSDVMSGLGILYMTYIGKTPFQO	736
QY	721	IINQISKLHAIIDPNHIE	PPDIPEKDQVLKCLKRDPKQKISIPPELLAHPPYQIQTH	780
Db	737	IINQISKLHAIIDPNHIE	PPDIPEKDQVLKCLKRDPKQKISIPPELLAHPPYQIQTH	796
QY	781	PVNQMAKGTTEEMKYV	LQVLGNSPNSILKAAKTLIYHSGESHSSSXTFFEKRGK	840
Db	797	PVNQMAKGTTEEMKYV	LQVLGNSPNSILKAAKTLIYHSGESHSSSXTFFEKRGK	856
QY	841	K 841		
Db	857	K 857		
RESULT 2				
Q8BY97	ID	Q8BY97	PRELIMINARY;	PRT; 815 AA.
AC	Q8BY97			
DT	01-MAR-2003	(TREMBlrel. 23, Created)		
DT	01-MAR-2003	(TREMBlrel. 23, Last sequence update)		
DT	01-OCT-2003	(TREMBlrel. 25, Last annotation update)		
DE		Dial specificity protein kinase TTK.		
GN	TKK.			
OS	Mus musculus (Mouse)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCEI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=Thymus;
RX	MEDLINE=22354683; PubMed=12466851;
RA	The FANTOM Consortium,
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;
RT	"Analysis of the mouse transcriptome based on functional annotation of
RT	60,770 full-length cDNAs.";
RL	Nature 420:563-573(2002).
DR	EMBL; AK041487; BAC30960.1; -.
DR	MGD; MGI:1194921; Ttk
DR	GO; GO:0005524; F:ATP binding; IEA.
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR	GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR	InterPro; IPR000719; Prot kinase.
DR	InterPro; IPR002290; Ser Thr_pkinase.
DR	InterPro; IPR008271; Ser Thr_pkin_AS.
DR	InterPro; IPR008941; Tyr-like.
DR	InterPro; IPR001245; Tyr_pkinase.
DR	Pfam; PF00069; Pkinase; 1.
DR	ProDom; PD000001; Prot kinase; 1.
DR	SMART; SM00219; Ty_Kc; 1.
DR	SMART; SM00219; Ty_Kc; 1.
DR	PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ	SEQUENCE 815 AA; 91524 MW; F12A2C0582B36767 CRC64;
Query Match 70.7%; Score 3092.5; DB 11; Length 815;	
Best Local Similarity 74.3%; Pred. No. 7.9e-189;	
Matches 604; Conservative 83; Mismatches 100; Indels 27; Gaps 7;	
QY	1 MNKVRDIKNFKNEDELTDLSLNKISADTTDNGSTVQIMMANNPEDWLSLLKLEKNS 60
Db	17 MSKTRDIKNKI-NEDCTDELSKICADHTE--TVNQIMRVGNTPENWLNFLKLEKNS 72
QY	61 VPLSDALLNLKIGRYSOAIEALPPDKYQONESFARIOVPAELKATQEPDDARDYFQMAR 120
Db	73 SPLNDLNLKIGRYSOAIEVLPPDKYQONESFARIOVRLAELKATQEPDDARDYFQMAR 132
QY	121 ANCKFAFVHISFAQFELSGCNVKSQQLLOKAVERGAVPLEMLEIALRNLNLOKKOLLS 180
Db	133 ENCKFAFVHVSFAQFELSGCNLKKSEQLLHKAVETGAVPLQMLETAMRNHLHKKQLLP 192
QY	181 EEBKXNLASVTIATAESFSGSLGHLQNRNNSCDSRGQTTKARFLYGENMPPQDAEIGYR 240
Db	193 EEDKXSVASVTLSAQEPFPSSSLGNVQNRISCSERQAGAAARVLYGENLPPQDAEVRHQ 252
QY	241 NSLRQTNKTKQSCPFGRVPVNLLNSPDCDVTKDSDVVPFCMRKQTSRSECRDLVVPGSKP 300
Db	253 NPFQTHAAKRSCPFGRVPVNLLNSPDFYVKTDSAVTQLTTRKSGPD-RDAILPGSRP 311
QY	301 SGNDSCELRLNLSQVNSHFKEPLVSDKSESLIITDSITLKNKTESLILAKLETKEQE 360
Db	312 RGSDSYELRGLKPIQTIYLKDSLVSNEKSEL-MSDLIALKSKTDS--LTKLEETK---- 365
QY	361 PEVPESNQKQWAKRSECIQNQNPAASSNHQIPELARKVNTQKHITTFEQPVFSVSKQS 420
Db	366 PEIAERPRMQWQSTRKDECVFQNPAPAPALRHVPDVTPKAD-----KES 409
QY	421 PPISTSKWFPKSI CKTPSSNTLDDYMSCFRTPVWKNDFPPACQLSTPYGQACFQOQOH 480
Db	410 PPISVPKWLDPK SACTPSSSSLDYMKCFKTPVWKNDFPPACPSSTPYSLARLQOQOQ 469
QY	481 QILATPLQNLQVLASSANECISVKGRIVSYILKQIGSGSSKVFOVLNKKQIYAIKYVN 540
Db	470 QGLSTPLQSLQISGSSSINECISVNGRIYSILKQIGSGSSKVFOVLNKKQINAIKYVN 529
QY	541 LEEADNQTLSYRNEIAYLNKLOQSDKIIRLYDYBITDQYIYMWMECGNIDLSNLWKKK 600
Db	530 LEDADSTIETSYRNEIAYLNKLOQSDKIIRLYDYBITDQYIYMWMECGNIDLSNLWKKK 589

```

Qy 601 KSIDPWKRSYKWMLEAVHTIHOHGIIVHSDLKPFANFLIVDGMKLIDFGIANOMOPDIT 660
Db 590 KSINPWKRSYKWMLEAVHTIHOHGIIVHSDLKPFANFLIVDGMKLIDFGIANOMOPDIT 649
Qy 661 SVKDSQVGTVMYMPPEAIKMDSSRRNGSKSKISPKSDVWSLGCILYMTYGTPTFQ 720
Db 650 SVKDSQVGTVMYMAPEAIRDMSSRRNSKIRTKVSPRSDVWSLGCILYMTYGTPTFQ 709
Qy 721 IINQISKLHALIDNHHIEPDPPEKDLQVLCCKLRDPKQISIPPELLAHFYVQIOH 780
Db 710 IINQVSKLHALIINPAHIEPPEISEKDLRDVLKCLVRNPKRISIPPELLTHFYVQIOH 769
Qy 781 PVNQMAKGTTEEMKYVLGQVLGNLSPNSILKA 814
Db 770 PGSQARGATDEMKYVLGQVLGNLSPNSILKTA 803

RESULT 3
Q98T92 PRELIMINARY; PRT; 882 AA.
AC Q98T92
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mpel/TKK.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Magnaghi-Jaulin L., Vigneron S., Lorca T., Labbe J.C.;
RT "X-Mpel/TKK is a novel Xenopus protein required for the spindle
RT checkpoint.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF250290; AAK27843.1; -
DR HSP; Q00534; IBI8.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006458; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PSS0011; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 882 AA; 99078 MW; 51C0F1B54DFDCD00 CRC64;

Query Match
Best Local Similarity 51.7%; Pred. No. 3.9e-129;
Matches 459; Conservative 134; Mismatches 208; Indels 87; Gaps 22;

Qy 1 MNKVRDINKFKNEE-LTDELSLNKISADTTDSNGTVNQIMMANNPDLWLSLLKLEKN 59
Db 17 LDRVKSPKTYGTDNDWDELTFSSGSDTTEHSGIFTH-LVTTKTPEEWLNCCLKLENT 75
Qy 60 SVPLSD-ALLNKLIGRSQAIEALPPDKYQNSFARIQVRAELKAIQEPDARDYFQM 118
Db 76 GLFOIDQLNKLIDNYSQAVGALPAEKSHNSYAKILVHFAELKAHLDDEARQFQF 135
Qy 119 ARANCKKFAFVHISFAQFELSQGNVKKSKOLLOKAVERGAVPLEMLEIALRLNLKQQL 178
Db 136 ARLNCKKFAFVHTAFQAFELSEGFKKCKQILRRGLECGAVPCPEMLNALKNVQFKPKQL 195
Qy 179 LSEEEKXNL--SASTVLTAQESFSG-SLGHQLQNRNNSCDRSGGQTTKARFLYGNM-PPQD 234

```

```

Db 196 ISDEKENLAVSSSHINOQTASFQNALGNPQ-RMKIESPEEYSVKTFYSGEKLSSEP 254
Qy 235 AE-IGYRNSLRQTNKTKOSCFGRVPVNLNSPCDV-KTDDSVVPCFMKQTSRSE--- 289
Db 255 FEDIGRKELLNMSAKT---CPLGRVPVQFATSPDTRTKSDGSGSSSVKRPFSMRVPV 311
Qy 290 -----CRDLVVPQS-----KPSGNDSCELRNLSKVQNSHFKEPLVSD 327
Db 312 LATLSNPXYLGDDLHCSEDIKVPSSNIIIPHEEQANEDSLDMKTPSSVILS-VTESNLIVK 370
Qy 328 KSESLIITDSITLKNKTESLILAKLETK--EYGEPEVPESNQKQWAKRSECEINQNP 385
Db 371 RNEDLIL-----GSTSTAVLNPOENSKPSESAPLLPNIN-----ARLSDTTKNNI 417
Qy 386 ASSNHWQIPELARKVNTQKHHTPE-OPVFSVSKQ-----SPPISTSKWFD 430
Db 418 EAEWKVKVETPRQI-----FQPMIFPAGKRKSTEPSANPGSRVRVPAPSLSKCD 469
Qy 431 PKSICKTFSSNTLDDYMSCFRTPVYKNDFFPACQLSTPYGQFACQOQHQILATPLQN- 489
Db 470 PVFVCGTTPVKSQEDYMKCFRTPVVKTNLGLAQMSTPYNNRSDYQOPQ-----TPVGQP 524
Qy 490 --LOVLASSANECISVKGRISYILKQISGSGSKVFOVLAKEKKIYAIKYNLEADNQ 547
Db 525 DCFPIASAFSSDCIIVKRAYAVLKQIGTSGSSKVQVMDDKHLYAIKYNLEADQ 584
Qy 548 TLDYSYRNIAVINKLQOHSKIRLYDYEITDQIYMYWMECGNIDLNLWKKKSIDPWE 607
Db 595 TIESYQNEISHLNKLOQHCXKILRFDYEITDQIYMYWMECGNIDLNLWKKKTIINWE 644
Qy 608 RKSYYWKNLEAVHTIHOHGIIVHSDLKPFANFLIVDGMKLIDFGIANOMOPDITSVKDSQ 667
Db 645 RKSYYWKNLEAVHTIHOHGIIVHSDLKPFANFLIVDGMKLIDFGIANOMOPDITSVKDSQ 704
Qy 668 VGTVMYMPPEAIKMDSSRRNGSKSKISPKSDVWSLGCILYMTYGTPTFQIINQISK 727
Db 705 VGTVMYMPPEAIKMDSSRRNGSKSKISPKSDVWSLGCILYMTYGTPTFQIINQIAK 764
Qy 728 LHAIDPNHEIEFPDPEKDLQVLCCKLRDPKQISIPPELLAHFYVQIOH--PVNQ 785
Db 765 LHSILDPGYEIEFPNPEKDLQVLCCKLRDPKQISIPPELLAHFYVQIOHPTQDOQV 824
Qy 786 AKGTTEEMKYVLGQVLGNLSPNSILKAAKTIYEHYSGESHSNSSSKT 833
Db 825 QKETTEEMKRIQLGLIGLNSPNSISRAAKNLDYQFNSGRSLDLSTLGT 872

RESULT 4
Q98X21 PRELIMINARY; PRT; 585 AA.
AC Q98X21
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Dual specificity protein kinase TTK (Fragment).
GN TTK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK049200; BAC33605.1; -
DR MGD; MGI:1194921; Ttk.
DR GO; GO:0005524; F:ATP binding; IEA.

```

DR	GO; GO:0004672; F:protein kinase activity; IEA.	RT	for zebrafish fin regeneration.;
DR	GO; GO:0004680; P:protein amino acid phosphorylation; IEA.	RL	Development 129:5141-5149 (2002).
DR	InterPro; IPR000719; Prot kinase.	DR	EMBL; AF488735; AAN61408.1; -
DR	InterPro; IPR008941; TPR-like.	DR	GO; GO:0005524; F:ATP binding; IEA.
DR	Pfam; PF00069; pkinase; 1.	DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR	ProDom; PD000001; Prot_kinase; 1.	DR	GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.	DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
FT	NON TER 585	DR	InterPro; IPR000719; Prot_kinase.
SQ	SEQUENCE 585 AA; 65550 MW; 26E683DFA6D561CF CRC64;	DR	InterPro; IPR002290; Ser_thr_kinase.
		DR	InterPro; IPR008271; TPR-like.
		DR	InterPro; IPR001245; Tyr_kinase.
		DR	Pfam; PF00069; pkinase; 1.
		DR	ProDom; PD000001; Prot_kinase; 1.
		DR	SMART; SM00220; S_TKc; 1.
		DR	SMART; SM00219; TYRKc; 1.
		DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
		DR	PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
		DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
		SQ	SEQUENCE 983 AA; 110087 MW; 4FB55FEAE1F10AFE CRC64;
		Query Match	37.7%; Score 1648.5; DB 13; Length 983;
		Best Local Similarity	39.4%; Pred.No.1.3e-96;
		Matches 400; Conservative	136; Mismatches 239; Indels 239; Gaps 28;
Qy	16 LTDELSLNK--ISADTTDNGTGNQIMMANNPEDMLSLLLKLEKNSVPLSD-ALLNKLI 72	Qy	16 LTDELSLNK--ISADTTDNGTGNQIMMANNPEDMLSLLLKLEKNSVPLSD-ALLNKLI 72
Db	16 LCQKAMKQLFNEDDTD--YINQ-AISSNSPTCTFTLSNLEKKNPOADPSLLSKLM 71	Db	16 LCQKAMKQLFNEDDTD--YINQ-AISSNSPTCTFTLSNLEKKNPOADPSLLSKLM 71
Qy	73 GRYSQAIEALPPDKYQNESFARIQVFAELKAIQEPDDARDYFQMARANCKYAFVHIS 132	Qy	73 GRYSQAIEALPPDKYQNESFARIQVFAELKAIQEPDDARDYFQMARANCKYAFVHIS 132
Db	72 DSYTRVFSMPLGKYSQESYAKMLVFAELKAIQVNDQAQTSFDIRSCKDPFVHVA 131	Db	72 DSYTRVFSMPLGKYSQESYAKMLVFAELKAIQVNDQAQTSFDIRSCKDPFVHVA 131
Qy	133 PAQFELSGNVKSKQLLOKAVRGAVPLEMLALRLNLQKQLLSEBEKNLSASTV 192	Qy	133 PAQFELSGNVKSKQLLOKAVRGAVPLEMLALRLNLQKQLLSEBEKNLSASTV 192
Db	132 YAQFELQGNKKCTMLQKAFEMNAKPRHVLAAVRLNLTGRQLLSHEDKENLSV-- 189	Db	132 YAQFELQGNKKCTMLQKAFEMNAKPRHVLAAVRLNLTGRQLLSHEDKENLSV-- 189
Qy	193 LTAQESFGSLHQLQNRNS---CDSRGQTTKARFLYG-----ENMP----- 231	Qy	193 LTAQESFGSLHQLQNRNS---CDSRGQTTKARFLYG-----ENMP----- 231
Db	190 -----ALDHQGGRRSDGTCELKPSST---FLHSSDQKTFSPQENGFWRTGSQHR 237	Db	190 -----ALDHQGGRRSDGTCELKPSST---FLHSSDQKTFSPQENGFWRTGSQHR 237
Qy	232 -----PQDAIGYRNS---LRQTN-----XTKOS 252	Qy	232 -----PQDAIGYRNS---LRQTN-----XTKOS 252
Db	238 RTAVARVPMVPLSIPENETSDSCAQKAEAPFTHSGFSRQTSRGSVRAFLSCSKKG 297	Db	238 RTAVARVPMVPLSIPENETSDSCAQKAEAPFTHSGFSRQTSRGSVRAFLSCSKKG 297
Qy	253 CPFG-----RVPVNLNLPDC---DVKTDDSVVPCF---MKRQTSRSECRDLVWPGSK 299	Qy	253 CPFG-----RVPVNLNLPDC---DVKTDDSVVPCF---MKRQTSRSECRDLVWPGSK 299
Db	298 TPDGDSYSLNKPVP---ISPDYLRDIEBEGHITALLNRAEKRETARTETTDINQIIS 354	Db	298 TPDGDSYSLNKPVP---ISPDYLRDIEBEGHITALLNRAEKRETARTETTDINQIIS 354
Qy	300 PSGNDSCE--LRNKK-----	Qy	300 PSGNDSCE--LRNKK-----
Db	355 TNSTEGQAFLKLNLEKADPHSDAAFLSKLLDCYSKYVFARFPLAEHCKTESYARMLVRYA 414	Db	355 TNSTEGQAFLKLNLEKADPHSDAAFLSKLLDCYSKYVFARFPLAEHCKTESYARMLVRYA 414
Qy	313 -----SVQNSHFK-----EPLVSDKSS 330	Qy	313 -----SVQNSHFK-----EPLVSDKSS 330
Db	415 ELKGIEDPDADDFSIARSHCKAFVHIAHAQFELSRGNSKSVSILQKALSSNARPI 474	Db	415 ELKGIEDPDADDFSIARSHCKAFVHIAHAQFELSRGNSKSVSILQKALSSNARPI 474
Qy	331 ELIITDSITLKNKTESLLAKLEETKEYQBPFE-----VPESKQKQWAKRKS 377	Qy	331 ELIITDSITLKNKTESLLAKLEETKEYQBPFE-----VPESKQKQWAKRKS 377
Db	475 ELLQTATRLNLSKGTLLLPADLQESSEAEVNAQNYIKRENSAKAPEDHQKPFKSTSS 534	Db	475 ELLQTATRLNLSKGTLLLPADLQESSEAEVNAQNYIKRENSAKAPEDHQKPFKSTSS 534
Qy	378 ECINQNPAAASHHWOIPELARKVNTQKHITTFEOPVFSV-----KQSPISSTSKWDPK 432	Qy	378 ECINQNPAAASHHWOIPELARKVNTQKHITTFEOPVFSV-----KQSPISSTSKWDPK 432
Db	535 E-----WKIPALITKHTSPEDRKASLEPVSSSSSHHVRTPAPLR-----LNPS 578	Db	535 E-----WKIPALITKHTSPEDRKASLEPVSSSSSHHVRTPAPLR-----LNPS 578
Qy	433 SICKTPSSNTLDDYMSCFRTPVKN-----DFPPACQLSTPYGQACFQQQHQIILATPL 487	Qy	433 SICKTPSSNTLDDYMSCFRTPVKN-----DFPPACQLSTPYGQACFQQQHQIILATPL 487
Db	579 LSCQTE--NYRQPNPNSGFVTPVVKQRPVIVSVAPATAQKMCPTALP-CTPQSGSVYIQPP- 634	Db	579 LSCQTE--NYRQPNPNSGFVTPVVKQRPVIVSVAPATAQKMCPTALP-CTPQSGSVYIQPP- 634
Qy	488 QNLQVLASSANECISVKGRISYILKQIGSGGSKVQVLNKKQIYAIKYVNLLEADNQ 547	Qy	488 QNLQVLASSANECISVKGRISYILKQIGSGGSKVQVLNKKQIYAIKYVNLLEADNQ 547
Db	635 --TQTPSFAFNSITINGQKQFFIKRMIGRGGSGSKVQVDFHKKHGVYAVKYVNLLEADAQ 692	Db	635 --TQTPSFAFNSITINGQKQFFIKRMIGRGGSGSKVQVDFHKKHGVYAVKYVNLLEADAQ 692

RESULT 5
Q8AYG3 PRELIMINARY; PRT; 983 AA.
AC Q8AYG3
AD Q8AYG3
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12399306;
RA Poss K.D., Neshiporuk A., Hillam A.M., Johnson S.L., Keating M.T.;
RT "Mps1 defines a proximal blastemal proliferative compartment essential

01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 Phosphotyrosine picked threonine kinase (PVT) (Fragment).
 Homo sapiens (Human).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 (1)
 SEQUENCE FROM N.A.
 MEDLINE=9314956; PubMed=7679926;
 Lindberg R.A., Fischer W.H., Hunter T.;
 "Characterization of a human protein threonine kinase isolated by
 screening an expression library with antibodies to phosphotyrosine";
 Oncogene 8:351-359(1993).
 -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; X70500; CAA49912.1; -;
 DR PIR; A42861;
 DR GO; GO:0005524; P:ATP binding; IEA.
 DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; P:transferase activity; IEA.
 DR GO; GO:0006458; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR002290; Ser Thr_pkinase.
 DR InterPro; IPR008271; Ser Thr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SMO0220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 FT NON_TER 305 305
 SQ SEQUENCE 305 AA; 34971 MW; 2B3F757EBCICE7D1 CRC64;
 Query Match 35.7%; Score 1564; DB 4; Length 305;
 Best Local Similarity 99.7%; Pred. No. 6.7e-92;
 Matches 298; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 488 QNLQVLAASSANECISVKGRYISLTQIGSGGSKYFOVLNERKQIYAIKYVNLNLEADNQ 547
 DB 5 QNLQVLAASSANECISVKGRYISLTQIGSGGSKYFOVLNERKQIYAIKYVNLNLEADNQ 64
 QY 548 TLDYSYRNEIAYLNKLOQHSDKIIRLYDYETDQIYIYMWECGNIDLSNMLKXKSIDPWE 607
 DB 65 TLDYSYRNEIAYLNKLOQHSDKIIRLYDYETDQIYIYMWECGNIDLSNMLKXKSIDPWE 124
 QY 608 RKSVMKMKLEAVHTIHOQGIYVHSDLPANFLIYDGMKLIIDFGIANQMOPDTTSVYKDSQ 667
 DB 125 RKSVMKMKLEAVHTIHOQGIYVHSDLPANFLIYDGMKLIIDFGIANQMOPDTTSVYKDSQ 184
 QY 668 VGTVMYVPPAIKDMSSRENGSKSKISPKSDVWSLGCILYIYMTYKTPFQOIINOISK 727
 DB 185 VGTVMYVPPAIKDMSSRENGSKSKISPKSDVWSLGCILYIYMTYKTPFQOIINOISK 244
 QY 728 LHAIDNHNIEFPDPEKDLQVLCCLKRPQKQISIPPELLAHYVQIOTHPVQMA 786
 DB 245 LHAIDNHNIEFPDPEKDLQVLCCLKRPQKQISIPPELLAHYVQIOTHPVQMA 303
 RESULT 8
 Q84VX4 PRELIMINARY; PRT; 777 AA.
 ID Q84VX4
 AC Q84VX4;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Atlg77720.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
 Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
 Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
 Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
 Palm C.C., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
 Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
 "Arabidopsis ORF clones";
 Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BT0046136; AAO42882.1; -;
 DR GO; GO:0005534; P:ATP binding; IEA.
 DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0006458; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR002290; Ser Thr_pkinase.
 DR InterPro; IPR008271; Ser Thr_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SMO0220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 SQ SEQUENCE 777 AA; 86376 MW; F2C25DA9609A878F CRC64;
 Query Match 14.8%; Score 647.5; DB 10; Length 777;
 Best Local Similarity 27.1%; Pred. No. 7e-33;
 Matches 217; Conservative 111; Mismatches 279; Indels 195; Gaps 26;
 QY 46 PEDWLSLLKLEKNSVPLSDALLNKLIGRYSOA-----IEALPPKYQGNSEFARIQV 98
 DB 15 PKSLVRPILNLETTSSSSSSSSPELRLHQAQKRLPSLKMQTTSIGPRKSVAPQRQ 74
 QY 99 RFAELKAT-----QBPDDARDYFQMARAN-----CKKFAFVHISFAOFLSQGNVKKSKQ 148
 DB 75 ASRNLRLVTAAGQSQDVVLTLSQSLAANTLTQDTNNLAITSVAGESASITPTVSEHFN 134
 QY 149 LLQKAVERGAVPLEMLEIALNLNKLQKLLSEEEKNLASVLTIAQE-----SFGSL 203
 DB 135 PSDRQMDFGKSAVTSLE---SNLDVORK-----SQSLIGTSQDMWDATNQAEA 180
 QY 204 GHLQNRNNSC--DSRGQTTKARFLYGENMPPQDAEIGVNSLSROTNK--TKQSCPFGRVPVN 261
 DB 181 SHL-----DAGTIGSKHQ-----NLPSVDSEVLSKSEYKSSSLAKIQGQLGEPN 225
 QY 262 LLNSPDCDVKTDDSVVPCFMKQTSRECDRLVVPVGSKPSGNDSCLELNLKSVQNSHFKE 321
 DB 226 FLNQP-----RTRC-----S 235
 QY 322 PLVDEKSESLIITDSITLKNKTBSLLAKLEETKEYOEPEVPSNOKQWAKRSECN 381
 DB 236 AVGSWATTTLIHSSAPMLNAT-----THVSRSYVEAD-----SN 271
 QY 382 QNPAAASNHWQIPELARKVNTTEQKHTTFEQPFVSQSPPISTSKWFDPKSICKTPSSN 441
 DB 272 ANPHAVQSQGNLPSCC---PSSKVSNIHLHPNKDASEMPASTN---DPEVRKEDTDS 324
 QY 442 TLDDYMSCFRTPVYVNDFFPACQLSTPYGPQ-----ACFQQQOHCILATPLQNL---QVLA 494
 DB 325 KQQQITTTGLEAPVG-----SSYSGDGGQANARLPBELHTSVSSQPKSKDKHEKVA 374
 QY 495 SSSANEC-----ISVKGRYISLTQIGSGGSKYFOVLNERKQIYAIKYVNLNLE 543
 DB 375 SSKGPSAPRKENYDPDLFFKVGKGLYQRLKIGSGGSEVHKVSISSDCTIYALKKIKLG 434
 QY 544 ADNOTLSYRNEIAYLNKLOQHSDKIIRLYDYETDQIYIYMWECGNIDLSNMLKXKSIDPWE 607
 DB 435 RYATATYGFQOEIGYLLKLGKTN-IIQLIDYEVTLLQEVINGTMSNKGGRVXEDGF 493
 QY 582 IYVMWECGNIDLSNMLKXK-----KSIDPWERKSYKMKMLEAVHTIHOQGIYVHSDLK 633

Db 494 IYMWLEGEIDLAHMLSKQKREIEGSDRTIDENWLRFYQOILQAVNTIHEERIVHSCLK 553
QY 634 PANFLIIVDGMKLIIDFGIANOMQDPTTSVVKDSQVGTWYVMPPEAIKDMSSSRENGSKS 693
Db 554 PANFLIVRGFLKLIIDFGIAKAINSDDTNIQRDSQVGTLSYMSPEAFM-CNESDENGNT-I 611
QY 694 KISPKSDVWSIGLCILYMYTKTFFQOIQINQISKLHAIIDPNHEIEFPDPEKDLQVCLK 753
Db 612 KCGRPSDIWSIGLCILYMYTKTFFQOIQINQISKLHAIIDPNHEIEFPDPEKDLQVCLK 671
QY 754 CCLKRDPEKORISIPPELLAHVYV 775
Db 672 KCLAWDRNQWRIPELLQHPFL 693

RESULT 9
Q9CA22
ID O9CA22 PRELIMINARY; PRT; 720 AA.
AC O9CA22;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T32E8.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA White O., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egtu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.B., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzali A.,
RA Mitscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana.";
RL Nature 408:816-820(2000).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AC012193; AAG51619.1; -.
DR F01; A96807; A96807.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00101; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
KW Hypothetical protein; ATP-binding; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 720 AA; 80027 MW; F26A9A0B1C6D91DE CRC64;

Query Match 14.7%; Score 643; DB 10; Length 720;

Best Local Similarity 28.1%; Pred. No. 1.2e-32;
Matches 205; Conservative 99; Mismatches 242; Indels 184; Gaps 24;
QY 107 QEPDARDYFQVARN-----CKFAFVHTISFAQFELSQGVNKKSKQLLOKAVRGAVP 160
Db 30 QRSQDVVLTSQSLANTLTQDTNNLAITSVAGESASITQPTVSEHPSPDRQMFQKSA 89
QY 161 LEMLEIALNLNLQKLLSEEEKNLGASTVLTAQE-----SFGSLGHLQNRNNSC-D 214
Db 90 VTSLE---SNLDVQRK-----SQSLIGTSQDMWDATNAQEAASHL-----DACIG 131
QY 215 SRGQTTKARPLYGENMPDQDAIGYRNSLRQTNK-TKQSCPPGRVPVNLNLSPCDCKVTD 273
Db 132 SKHQ-----NLPSVDSSEVLSKSEYKDSSSSLAKIQQLGEPF-NFLNQP----- 173
QY 274 DSVWPCFMKQTSRSECDRLVVPFGSKPQNSCELRNLKSVQNSHFKEPLVSDKSSSLI 333
Db 174 -----RTRC-----SAVGSWATTTLI 190
QY 334 ITDSITLKNKTESSILLAKLETKEYQBPPEVSNOKQWAKKSCINQNPAAASHHWQI 393
Db 191 HSSAPMLNAT-----THVRSRYEAD-----SNANHAVOSQGNL 226
QY 394 PELARKVNTQKHTTPEOPVPSVSKQSPPISTKWFDPKSICTPSSNTLDDYMSCFRTP 453
Db 227 PSCC-----PSSKVSNIILHFNKDATASEMPASTN---DPEVRVKETDTSKQOQITGLEAP 279
QY 454 VVKNDPPACQLSTPYGPQ---ACFQQOQHQILATPQLN---QVLASSSANEC----- 501
Db 280 VG-----SSIYGSQGANARLPBELHTSVSSQPKSKDEKVASSGKSPAPRKN 329
QY 502 -----ISVKGRISILKIQSGSGSKGVFVLNEKKQIYAIKYVNLNEADNQTLDSEYNE 555
Db 330 YPDJFFKVGKLYQRLKIGSGSGSEVHKVSISSCTIYALKKIKLGRDVATAYGFOQE 389
QY 556 IAYLNKLQHSKRIILYDYEITDQ-----YIYMWECGNIDL 593
Db 390 IGYLKLKGTN-IIQLIDYVTDKTLQELVNGTMSNKGDRVKEDGPIYVWLEYEIDL 448
QY 594 NSWLKK-----KSIDPWERKSKYKMWLEAVHTIHOHGIVHSDLKPAFLIYDGLMK 645
Db 449 AHMLSKQKREIEGSDRTIDENWLRFYQOILQAVNTIHEERIVHSDLKPAFLIYDGLMK 508
QY 646 LIDFGIANOMQDPTTSVVKDSQVGTWYVMPPEAIKDMSSSRENGSKSKISPKSDVWSIG 705
Db 509 LIDFGIAKAINSDDTNIQRDSQVGTLSYMSPEAFM-CNESDENGNT-IKCGRPSDIWSIG 566
QY 706 CILYMYTKTFFQOIQINQISKLHAIIDPNHEIEFPDPEKDLQVCLKRDKPKQKIS 765
Db 567 CILYQWYGRTPFADYKTFWAKFKVITDPNHEITYNQLSNPWLIDLMKKCLAWDRNQWR 626
QY 766 IPELLAHVYV 775
Db 627 IPELLQHPFL 636

RESULT 10
ID O8LN55 PRELIMINARY; PRT; 742 AA.
AC O8LN55;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OSJNBA0004G17.15.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;

Search completed: September 29, 2004, 16:20:43
Job time : 144 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 16:16:51 ; Search time 144 Seconds
(without alignments)
1879.397 Million cell updates/sec

Title: US-10-081-119-14

Perfect score: 4376

Sequence: 1 MNKVRDIKNKFKNEDLTDEL.....GGESHNSSSSKTEFKRGKK 841

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4376	100.0	841	14	US-10-081-119-14
2	4376	100.0	841	14	Sequence 14, Appl
3	4376	100.0	841	15	Sequence 666, Appl
4	4376	100.0	841	15	Sequence 2, Appl
5	4376	100.0	841	15	Sequence 16, Appl
6	4376	100.0	841	15	Sequence 804, Appl
7	4376	100.0	841	15	Sequence 833, Appl
8	4376	100.0	841	15	Sequence 865, Appl
9	4376	100.0	857	14	Sequence 30, Appl
10	4373	99.9	857	15	Sequence 482, Appl
11	1929	44.1	367	16	Sequence 34, Appl
12	673.5	15.4	678	14	Sequence 157, Appl
13	637.5	14.6	742	16	Sequence 32, Appl
14	627	14.3	730	12	Sequence 201272,
15	614	14.0	399	12	Sequence 273093,
					Sequence 43102, A

16	603	13.8	735	14	US-10-128-714-3040	Sequence 3040, Ap
17	603	13.8	842	14	US-10-128-714-8040	Sequence 8040, Ap
18	586.5	13.4	764	14	US-10-081-119-30	Sequence 30, Appl
19	586.5	13.4	764	15	US-10-369-493-1595	Sequence 1595, Ap
20	558	12.8	690	14	US-10-032-585-7510	Sequence 7510, Ap
21	553.5	12.6	399	15	US-10-369-493-3420	Sequence 3420, Ap
22	352.5	8.1	256	12	US-09-964-356-72	Sequence 72, Appl
23	352.5	8.1	256	12	US-10-042-865-83	Sequence 83, Appl
24	352.5	8.1	256	12	US-10-029-020-183	Sequence 183, App
25	352.5	8.1	256	12	US-10-072-012-799	Sequence 799, App
26	352.5	8.1	256	12	US-10-072-012-856	Sequence 856, App
27	352.5	8.1	256	12	US-10-072-012-876	Sequence 876, App
28	352.5	8.1	256	12	US-10-072-012-882	Sequence 882, App
29	352.5	8.1	256	16	US-10-055-569A-99	Sequence 99, Appl
30	349.5	8.0	254	15	US-10-074-978A-150	Sequence 150, Appl
31	345.5	7.9	255	12	US-10-087-684-98	Sequence 98, Appl
32	345.5	7.9	255	12	US-10-218-779-98	Sequence 98, Appl
33	344.5	7.9	256	10	US-09-823-187-96	Sequence 96, Appl
34	343.5	7.8	256	10	US-09-863-776-40	Sequence 40, Appl
35	342.5	7.8	749	9	US-09-771-161A-261	Sequence 261, Appl
36	340.5	7.8	251	15	US-10-052-648A-41	Sequence 41, Appl
37	338	7.7	528	14	US-10-032-585-7571	Sequence 7571, Ap
38	335.5	7.7	252	15	US-10-074-978A-160	Sequence 160, App
39	332.5	7.6	250	10	US-09-976-782-31	Sequence 31, Appl
40	329.5	7.5	749	12	US-10-092-900A-20	Sequence 20, Appl
41	328.5	7.5	250	12	US-10-042-865-102	Sequence 102, App
42	328.5	7.5	787	14	US-10-354-358-48	Sequence 48, Appl
43	328.5	7.5	787	16	US-10-408-765A-267	Sequence 267, Appl
44	327	7.5	460	9	US-09-740-627-10	Sequence 10, Appl
45	319	7.3	231	9	US-09-515-806-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-10-081-119-14
; Sequence 14, Application US/10081119
; Publication No. US20030045491A1
; GENERAL INFORMATION:
; APPLICANT: Reinhard, Christoph
; APPLICANT: Jefferson, Anne B.
; APPLICANT: Chan, Vivien W.
; TITLE OF INVENTION: TTK in Diagnosis and as a Therapeutic
; TITLE OF INVENTION: Target in Cancer
; FILE REFERENCE: 16932.002
; CURRENT APPLICATION NUMBER: US/10/081,119
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/289,813
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0.
; SEQ ID NO 14
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-081-119-14

Query Match	100.0%	Score 4376;	DB 14;	Length 841;
Best Local Similarity	100.0%	Pred. No. 2.4e-299;		
Matches 841;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MNKVRDIKNKFKNEDLTDELSLNKISADTTDSNGTVNQIMMANNPEOWLSLLKLEKNS	60	
Db	1	MNKVRDIKNKFKNEDLTDELSLNKISADTTDSNGTVNQIMMANNPEOWLSLLKLEKNS	60	
Qy	61	VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIQVRFALKAIOEPPDDARDYFQMAR	120	
Db	61	VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIQVRFALKAIOEPPDDARDYFQMAR	120	
Qy	121	ANCKKFAFVHISFAQFELSQGNVKKSKQLLQKAVERGAVPLEMLEIALNLNLOKQKLLS	180	
Db	121	ANCKKFAFVHISFAQFELSQGNVKKSKQLLQKAVERGAVPLEMLEIALNLNLOKQKLLS	180	

```
181 BEEKNLASVTLTAQSFSGSLGHLQNRNNSCDRSGQTTKARFLYGENMPPQDAEIGYR 240
181 BEEKNLASVTLTAQSFSGSLGHLQNRNNSCDRSGQTTKARFLYGENMPPQDAEIGYR 240
241 NSLRQTNKTKOSCPGPRVFNLLNSPCDVKTDDSVVPCFMKRGTSSECRDLVWPGSKP 300
241 NSLRQTNKTKOSCPGPRVFNLLNSPCDVKTDDSVVPCFMKRGTSSECRDLVWPGSKP 300
301 SGNDSCELNLKSVQNSHFKEPLVSDKSELIITDSITLKNKTESLLAKLEETKEYOE 360
301 SGNDSCELNLKSVQNSHFKEPLVSDKSELIITDSITLKNKTESLLAKLEETKEYOE 360
361 PEVPESKQKQAKRKSECCINONPAASNHWQIPELARKVNTQKHTTTPQPVFSVSKOS 420
361 PEVPESKQKQAKRKSECCINONPAASNHWQIPELARKVNTQKHTTTPQPVFSVSKOS 420
421 PPISTSKWFDPKSICKTPSSNTLDDYMSCRTPVVKNDFFPACQLSTPYGQACFQOQOH 480
421 PPISTSKWFDPKSICKTPSSNTLDDYMSCRTPVVKNDFFPACQLSTPYGQACFQOQOH 480
481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGGSKVFQVLNEKKQIYAIKYVN 540
481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGGSKVFQVLNEKKQIYAIKYVN 540
541 LEEADNQTLDYSRNEIAYLNKLOQHSKIIRLDYEITDQIYIYMWECGNIDLSNLKXK 600
541 LEEADNQTLDYSRNEIAYLNKLOQHSKIIRLDYEITDQIYIYMWECGNIDLSNLKXK 600
601 KSIDPWERKSWKXNMLEAVHTIHOHGIVHSDLPANFLIYDGMKLIIDFGIANQMPDIT 660
601 KSIDPWERKSWKXNMLEAVHTIHOHGIVHSDLPANFLIYDGMKLIIDFGIANQMPDIT 660
661 SVVKDSQVGTVNYMPPPAIKDMSSSRENGSKSKISPKSDVWSLGCILYIYMTYKTPFQ 720
661 SVVKDSQVGTVNYMPPPAIKDMSSSRENGSKSKISPKSDVWSLGCILYIYMTYKTPFQ 720
721 IINQISKLHAIIDPNHEIEFPDIPEKLDQVLKCLKRDPKQISIPELLAHPYVQIOTH 780
721 IINQISKLHAIIDPNHEIEFPDIPEKLDQVLKCLKRDPKQISIPELLAHPYVQIOTH 780
781 PVNQAKGTTTEEMKYVLGQVLGNLSPNSILKAATLYEHYSGGSHNSSSSKTTEKRGK 840
781 PVNQAKGTTTEEMKYVLGQVLGNLSPNSILKAATLYEHYSGGSHNSSSSKTTEKRGK 840
841 K 841
841 K 841
```

RESULT 2
US-10-116-712-666
; Sequence 666, Application US/10116712
; Publication No. US20030194764A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Switzer, Ann
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.568
; CURRENT APPLICATION NUMBER: US/10/116,712
; CURRENT FILING DATE: 2002-04-07
; NUMBER OF SEQ ID NOS: 670
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 666
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-712-666

Query Match 100.0%; Score 4376; DB 14; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.4e-299;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MNKVRDINKKFNEDLTDELNLNKISADTTNSGTVQIMMMANNPEDWLSLLKLEKNS 60
DB 1 MNKVRDINKKFNEDLTDELNLNKISADTTNSGTVQIMMMANNPEDWLSLLKLEKNS 60
QY 61 VPLSDALLNKILGRYSOAIEALPDKYQONESFARIQVRFAELKAIQEPDDARDYFQWAR 120
DB 61 VPLSDALLNKILGRYSOAIEALPDKYQONESFARIQVRFAELKAIQEPDDARDYFQWAR 120
121 ANCKKFAFVHISFAQFELSOGNVKSKOLLOKXAVERGAVPLEMLELIALNLNLOKKOLLS 180
121 ANCKKFAFVHISFAQFELSOGNVKSKOLLOKXAVERGAVPLEMLELIALNLNLOKKOLLS 180
181 BEEKNLASVTLTAQSFSGSLGHLQNRNNSCDRSGQTTKARFLYGENMPPQDAEIGYR 240
181 BEEKNLASVTLTAQSFSGSLGHLQNRNNSCDRSGQTTKARFLYGENMPPQDAEIGYR 240
241 NSLRQTNKTKOSCPGPRVFNLLNSPCDVKTDDSVVPCFMKRGTSSECRDLVWPGSKP 300
241 NSLRQTNKTKOSCPGPRVFNLLNSPCDVKTDDSVVPCFMKRGTSSECRDLVWPGSKP 300
301 SGNDSCELNLKSVQNSHFKEPLVSDKSELIITDSITLKNKTESLLAKLEETKEYOE 360
301 SGNDSCELNLKSVQNSHFKEPLVSDKSELIITDSITLKNKTESLLAKLEETKEYOE 360
361 PEVPESKQKQAKRKSECCINONPAASNHWQIPELARKVNTQKHTTTPQPVFSVSKOS 420
361 PEVPESKQKQAKRKSECCINONPAASNHWQIPELARKVNTQKHTTTPQPVFSVSKOS 420
421 PPISTSKWFDPKSICKTPSSNTLDDYMSCRTPVVKNDFFPACQLSTPYGQACFQOQOH 480
421 PPISTSKWFDPKSICKTPSSNTLDDYMSCRTPVVKNDFFPACQLSTPYGQACFQOQOH 480
481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGGSKVFQVLNEKKQIYAIKYVN 540
481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGGSKVFQVLNEKKQIYAIKYVN 540
541 LEEADNQTLDYSRNEIAYLNKLOQHSKIIRLDYEITDQIYIYMWECGNIDLSNLKXK 600
541 LEEADNQTLDYSRNEIAYLNKLOQHSKIIRLDYEITDQIYIYMWECGNIDLSNLKXK 600
601 KSIDPWERKSWKXNMLEAVHTIHOHGIVHSDLPANFLIYDGMKLIIDFGIANQMPDIT 660
601 KSIDPWERKSWKXNMLEAVHTIHOHGIVHSDLPANFLIYDGMKLIIDFGIANQMPDIT 660
661 SVVKDSQVGTVNYMPPPAIKDMSSSRENGSKSKISPKSDVWSLGCILYIYMTYKTPFQ 720
661 SVVKDSQVGTVNYMPPPAIKDMSSSRENGSKSKISPKSDVWSLGCILYIYMTYKTPFQ 720
721 IINQISKLHAIIDPNHEIEFPDIPEKLDQVLKCLKRDPKQISIPELLAHPYVQIOTH 780
721 IINQISKLHAIIDPNHEIEFPDIPEKLDQVLKCLKRDPKQISIPELLAHPYVQIOTH 780
781 PVNQAKGTTTEEMKYVLGQVLGNLSPNSILKAATLYEHYSGGSHNSSSSKTTEKRGK 840
781 PVNQAKGTTTEEMKYVLGQVLGNLSPNSILKAATLYEHYSGGSHNSSSSKTTEKRGK 840
841 K 841
841 K 841
```

RESULT 3
US-10-353-690-2
; Sequence 2, Application US/10353690
; Publication No. US20030215840A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Chun, Miyoung
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.
; APPLICANT: Donoghue, Mary

APPLICANT: Stagliano, Nancy
APPLICANT: Perodin, Jacqueline
APPLICANT: Rodrique-Way, Amelie
TITLE OF INVENTION: Methods and compositions for treating
TITLE OF INVENTION: cardiovascular disease using 1682, 6169, 6193, 7771, 14395,
TITLE OF INVENTION: 29002, 33216, 43726, 69292, 26156, 32427, 2402, 7747, 1720,
TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 9252, 9135,
TITLE OF INVENTION: 10532, 18610, 8165, 2445, 64624, 84237, 8912, 2868,
TITLE OF INVENTION: 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914,
TITLE OF INVENTION: 51130, 19489, 21833, 2917, 59590, 15592, 2094, 2252, 3474,
TITLE OF INVENTION: 9792, 15400, 1452 or 6585 molecules
FILE REFERENCE: MPI02-018P1RNMNM
CURRENT APPLICATION NUMBER: US/10/353,690
CURRENT FILING DATE: 2003-01-29
PRIOR APPLICATION NUMBER: 60/353,224
PRIOR FILING DATE: 2002-02-01
PRIOR APPLICATION NUMBER: 60/364,529
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: 60/373,861
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/376,287
PRIOR FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 60/389,080
PRIOR FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: 60/390,971
PRIOR FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: 60/394,130
PRIOR FILING DATE: 2002-07-03
PRIOR APPLICATION NUMBER: 60/394,797
PRIOR FILING DATE: 2002-07-10
PRIOR APPLICATION NUMBER: 60/404,904
PRIOR FILING DATE: 2002-08-21
PRIOR APPLICATION NUMBER: 60/405,450
PRIOR FILING DATE: 2002-08-23
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 126
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 841
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-353-690-2

Query Match 100.0%; Score 4376; DB 15; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.4e-299;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKVRDIKXFKNEDLTDELSLNKISADTTDNGTQVNMANNPEDWLSLLKLEKNS 60
DB 1 MNKVRDIKXFKNEDLTDELSLNKISADTTDNGTQVNMANNPEDWLSLLKLEKNS 60

QY 61 VPLSDALLNLKLGYSQAIEALPDPKYQNESFARIOVRFAELKATQEPDARDYFQMAR 120
DB 61 VPLSDALLNLKLGYSQAIEALPDPKYQNESFARIOVRFAELKATQEPDARDYFQMAR 120

QY 121 ANCKKFAFVHISFAQFELSQNVKKSQQLQKAVERGAVPLEMLETALRNLNKQKQLLS 180
DB 121 ANCKKFAFVHISFAQFELSQNVKKSQQLQKAVERGAVPLEMLETALRNLNKQKQLLS 180

QY 181 EEEKNNLSASTVLTAQSFSGSLGHLQNRNNSCDRQCTTKARFLYGENMPPQDAIGYR 240
DB 181 EEEKNNLSASTVLTAQSFSGSLGHLQNRNNSCDRQCTTKARFLYGENMPPQDAIGYR 240

QY 241 NSLRQTKTKQSCFPGRVPVNLNSPCDVKTDDSVVPCFMKQTSRSECRDLVPGSKP 300
DB 241 NSLRQTKTKQSCFPGRVPVNLNSPCDVKTDDSVVPCFMKQTSRSECRDLVPGSKP 300

QY 301 SGNDSCELRNKLSYQNSHFKEPLVSDKSSLELIITDSITLKNKTESLLAKLEETKEYQE 360
DB 301 SGNDSCELRNKLSYQNSHFKEPLVSDKSSLELIITDSITLKNKTESLLAKLEETKEYQE 360

QY 361 PEVPESNQKQWAKRSEKINQNPAASSNHQWIPELARKVNTQKHTTPEQVPFSVSKQS 420

DB 361 PEVPESNQKQWAKRSEKINQNPAASSNHQWIPELARKVNTQKHTTPEQVPFSVSKQS 420
QY 421 PRISTSKMFPDPKSIKCTPSSNTLDDYMSCFRTPVVKNDPPACQLSTPYGQACFQQOQH 480
DB 421 PRISTSKMFPDPKSIKCTPSSNTLDDYMSCFRTPVVKNDPPACQLSTPYGQACFQQOQH 480
QY 481 QILATPLQNLQVLASSANECISVKGRISYILKQIGSGGSSKVFQVLNEKKQIYAIKYVN 540
DB 481 QILATPLQNLQVLASSANECISVKGRISYILKQIGSGGSSKVFQVLNEKKQIYAIKYVN 540
QY 541 LEEADNQTLDYSRNEIAYLNKLQOHSKIIIRLYDEITDQYIYMYMECGNIDLSWLKXK 600
DB 541 LEEADNQTLDYSRNEIAYLNKLQOHSKIIIRLYDEITDQYIYMYMECGNIDLSWLKXK 600
QY 601 KSIDPWERKSYKWNMLEAVHTTHOIGIVHSDLKPANFLIVDGMKLKIDFGIANQMPDPT 660
DB 601 KSIDPWERKSYKWNMLEAVHTTHOIGIVHSDLKPANFLIVDGMKLKIDFGIANQMPDPT 660
QY 661 SVVKDSQVGTVNYMPPPEAIKDMSSSRENGKSKSIKSPKSDVMSLGCILYYTYGKTPFQ 720
DB 661 SVVKDSQVGTVNYMPPPEAIKDMSSSRENGKSKSIKSPKSDVMSLGCILYYTYGKTPFQ 720
QY 721 IINQISKLHAIIDPNHEIEFPDIPKDIQDVVKCCLKRDPKORISIPELLAPHYVQIOH 780
DB 721 IINQISKLHAIIDPNHEIEFPDIPKDIQDVVKCCLKRDPKORISIPELLAPHYVQIOH 780
QY 781 PYNQMAKGTTEEMKYVLGQVLGNSPNSILKAAKTLYHSGSGSHSSSSKTFEKGK 840
DB 781 PYNQMAKGTTEEMKYVLGQVLGNSPNSILKAAKTLYHSGSGSHSSSSKTFEKGK 840
QY 841 K 841
DB 841 K 841

RESULT 4
US-10-366-288-15
Sequence 16, Application US/10366288
Publication No. US20030216288A1
GENERAL INFORMATION:
APPLICANT: Powell, Douglas
APPLICANT: Weich, Nadine S.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
TITLE OF INVENTION: AIDS AND HIV-RELATED DISORDERS USING 1414, 1481, 1553,
TITLE OF INVENTION: 34021, 1720, 1683, 1552, 1682, 1675, 12825, 3952, 5816,
TITLE OF INVENTION: 10002, 1611, 1371, 14324, 126, 270, 312, 167, 18926,
TITLE OF INVENTION: 6747, 1793, 1784 OR 2045 MOLECULES
FILE REFERENCE: MPI02-025P1RNMNM
CURRENT APPLICATION NUMBER: US/10/366,288
CURRENT FILING DATE: 2003-02-13
PRIOR APPLICATION NUMBER: 60/357,391
PRIOR FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: 60/380,249
PRIOR FILING DATE: 2002-05-13
PRIOR APPLICATION NUMBER: 60/391,306
PRIOR FILING DATE: 2002-06-25
PRIOR APPLICATION NUMBER: 60/406,297
PRIOR FILING DATE: 2002-08-27
PRIOR APPLICATION NUMBER: 60/412,007
PRIOR FILING DATE: 2002-09-19
PRIOR APPLICATION NUMBER: 60/417,508
PRIOR FILING DATE: 2002-10-10
PRIOR APPLICATION NUMBER: 60/432,318
PRIOR FILING DATE: 2002-12-10
NUMBER OF SEQ ID NOS: 52
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 841
TYPE: PRT
ORGANISM: Homo Sapien
US-10-366-288-15

```

Query Match      100.0%; Score 4376; DB 15; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.4e-299;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNKVRDIKNKFKNEBTLDELSLNKISADTTDNGSTVNOIMMANNPEDWLSLLKLEKNS 60
Db 1 MNKVRDIKNKFKNEBTLDELSLNKISADTTDNGSTVNOIMMANNPEDWLSLLKLEKNS 60

Qy 61 VPLSDALLNLKIGRYSAIEALPPDKYQONESFARIQVRFPAELKAIQEPDARDYFQMAR 120
Db 61 VPLSDALLNLKIGRYSAIEALPPDKYQONESFARIQVRFPAELKAIQEPDARDYFQMAR 120

Qy 121 ANCKKFAFVHISFAQFELSQGNVKKSKQLQKAVERGAVPLEMLEIALRNINLQKKLLS 180
Db 121 ANCKKFAFVHISFAQFELSQGNVKKSKQLQKAVERGAVPLEMLEIALRNINLQKKLLS 180

Qy 181 EEEKKNLSASTVLTQAQSFSGSLGHLQNRNNSCDSRGQTTKARFLYGENNPPQDAEIGYR 240
Db 181 EEEKKNLSASTVLTQAQSFSGSLGHLQNRNNSCDSRGQTTKARFLYGENNPPQDAEIGYR 240

Qy 241 NSLRQTNKTKQSCPFGRVPVNLNSPCDVKTDDSVVPCFMKQTSRSECRDLVPGSKP 300
Db 241 NSLRQTNKTKQSCPFGRVPVNLNSPCDVKTDDSVVPCFMKQTSRSECRDLVPGSKP 300

Qy 301 SGNDSCELRNLSKVSQNSHFKEPLVSDKSSBELIITDSITLKNKTESLLAKLEETKEYQE 360
Db 301 SGNDSCELRNLSKVSQNSHFKEPLVSDKSSBELIITDSITLKNKTESLLAKLEETKEYQE 360

Qy 361 PEVPESQKQWQAKRKSECINQNPAASSNHQWQIPELARKVNTBQKHTTFEQPVFSVSKQS 420
Db 361 PEVPESQKQWQAKRKSECINQNPAASSNHQWQIPELARKVNTBQKHTTFEQPVFSVSKQS 420

Qy 421 PPISTSKWFDPKSCKTPSSNTLDDVNSCFRTPVKNDPPACOLSTPYGQPCACFOQOQH 480
Db 421 PPISTSKWFDPKSCKTPSSNTLDDVNSCFRTPVKNDPPACOLSTPYGQPCACFOQOQH 480

Qy 481 QILATPLQNLQVLASSANECISVKGRISYILKQIGSGSSKVFQVLNKKQIYAIKYVN 540
Db 481 QILATPLQNLQVLASSANECISVKGRISYILKQIGSGSSKVFQVLNKKQIYAIKYVN 540

Qy 541 LEEADNQTLSYRNEIAYLNKLOHSDKIIRLYDYEITDQIYIYVMECGNIDLSNLKXK 600
Db 541 LEEADNQTLSYRNEIAYLNKLOHSDKIIRLYDYEITDQIYIYVMECGNIDLSNLKXK 600

Qy 601 KSIDPWERKSWKKNLEAVHTIHOHGIHVSDLPANFLIVDGMKLIIDFGIANQMPDIT 660
Db 601 KSIDPWERKSWKKNLEAVHTIHOHGIHVSDLPANFLIVDGMKLIIDFGIANQMPDIT 660

Qy 661 SVVKDSQVGTNYNMPPEAIKDMSSSRENGSKSKISPKSDVWSLGLIYMTYKTPFQO 720
Db 661 SVVKDSQVGTNYNMPPEAIKDMSSSRENGSKSKISPKSDVWSLGLIYMTYKTPFQO 720

Qy 721 IINOISKLHAIIDPNHIEFPDPEKDLQVLCCLKRPDKORISIPELLAHPYVOIQTH 780
Db 721 IINOISKLHAIIDPNHIEFPDPEKDLQVLCCLKRPDKORISIPELLAHPYVOIQTH 780

Qy 781 PVNQMAXGTTTEEMKYVLQVLGNSPNSILKAAKTYEHYSGGSHNSSSSKTFEKKRGK 840
Db 781 PVNQMAXGTTTEEMKYVLQVLGNSPNSILKAAKTYEHYSGGSHNSSSSKTFEKKRGK 840

841 K 841
841 K 841

```

RESULT 5

US-10-295-027-804
; Sequence 804 Application US/10295027
; Publication No. US2003023250A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.

```

; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 804
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-804

```

```

Query Match      100.0%; Score 4376; DB 15; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.4e-299;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNKVRDIKNKFKNEBTLDELSLNKISADTTDNGSTVNOIMMANNPEDWLSLLKLEKNS 60
Db 1 MNKVRDIKNKFKNEBTLDELSLNKISADTTDNGSTVNOIMMANNPEDWLSLLKLEKNS 60

Qy 61 VPLSDALLNLKIGRYSAIEALPPDKYQONESFARIQVRFPAELKAIQEPDARDYFQMAR 120
Db 61 VPLSDALLNLKIGRYSAIEALPPDKYQONESFARIQVRFPAELKAIQEPDARDYFQMAR 120

Qy 121 ANCKKFAFVHISFAQFELSQGNVKKSKQLQKAVERGAVPLEMLEIALRNINLQKKLLS 180
Db 121 ANCKKFAFVHISFAQFELSQGNVKKSKQLQKAVERGAVPLEMLEIALRNINLQKKLLS 180

Qy 181 EEEKKNLSASTVLTQAQSFSGSLGHLQNRNNSCDSRGQTTKARFLYGENNPPQDAEIGYR 240
Db 181 EEEKKNLSASTVLTQAQSFSGSLGHLQNRNNSCDSRGQTTKARFLYGENNPPQDAEIGYR 240

Qy 241 NSLRQTNKTKQSCPFGRVPVNLNSPCDVKTDDSVVPCFMKQTSRSECRDLVPGSKP 300
Db 241 NSLRQTNKTKQSCPFGRVPVNLNSPCDVKTDDSVVPCFMKQTSRSECRDLVPGSKP 300

Qy 301 SGNDSCELRNLSKVSQNSHFKEPLVSDKSSBELIITDSITLKNKTESLLAKLEETKEYQE 360
Db 301 SGNDSCELRNLSKVSQNSHFKEPLVSDKSSBELIITDSITLKNKTESLLAKLEETKEYQE 360

Qy 361 PEVPESQKQWQAKRKSECINQNPAASSNHQWQIPELARKVNTBQKHTTFEQPVFSVSKQS 420
Db 361 PEVPESQKQWQAKRKSECINQNPAASSNHQWQIPELARKVNTBQKHTTFEQPVFSVSKQS 420

```

QY 421 PPISTKWFDPKSICTKTPSSNTLDDYMSCFRTPVVKNDFPPACOLSTPYGPACFOQOQH 480
Db 421 PPISTKWFDPKSICTKTPSSNTLDDYMSCFRTPVVKNDFPPACOLSTPYGPACFOQOQH 480
QY 481 QILATPLQNLQVLASSANECISVKGRYISILKQIGSGSKVQVNLNKKQIYAIKYN 540
Db 481 QILATPLQNLQVLASSANECISVKGRYISILKQIGSGSKVQVNLNKKQIYAIKYN 540
QY 541 LEEADNQTLDYSYRNEIAYLNKLQOHSDKIIRLYDYETDQIYVMWECGNIDLSWLKXK 600
Db 541 LEEADNQTLDYSYRNEIAYLNKLQOHSDKIIRLYDYETDQIYVMWECGNIDLSWLKXK 600
QY 601 KSIDPWERKSYWKNMLEAVHTIHQGHVHSDLPANFLIVDGMKLIDFGIANQMPDPT 660
Db 601 KSIDPWERKSYWKNMLEAVHTIHQGHVHSDLPANFLIVDGMKLIDFGIANQMPDPT 660
QY 661 SVVKDSQVGTVMYMPPEAIKDMSSRENGSKSKISPKSDVWSLGCILYMTYKTPFQ 720
Db 661 SVVKDSQVGTVMYMPPEAIKDMSSRENGSKSKISPKSDVWSLGCILYMTYKTPFQ 720
QY 721 IINOISKLHAIIDPNHEIEFPDIPEKQDVLKCLKRPKQIRISIPELLAHPYVQIOTH 780
Db 721 IINOISKLHAIIDPNHEIEFPDIPEKQDVLKCLKRPKQIRISIPELLAHPYVQIOTH 780
QY 781 PVNOMAKGTTTEEMKYVGLQVGLNSPNSILKAAKTLYEHYSGGSHNSSSKTFEKKRGK 840
Db 781 PVNOMAKGTTTEEMKYVGLQVGLNSPNSILKAAKTLYEHYSGGSHNSSSKTFEKKRGK 840
QY 841 K 841
Db 841 K 841

RESULT 6
US-10-295-027-833
; Sequence 833, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glyne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 833
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-833

Query Match 100.0%; Score 4376; DB 15; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.4e-299; Indels 0; Gaps 0;
Matches 841; Conservative 0; Mismatches 0;
QY 1 MNKVRDIKKNFKNEDLTDELSLNKISADTTDNGSVTNQIMMANNPDEWLSLLKLEKNS 60
Db 1 MNKVRDIKKNFKNEDLTDELSLNKISADTTDNGSVTNQIMMANNPDEWLSLLKLEKNS 60
QY 61 VPLSDALLNKLGRYSQAIEALPPDKYQNESFAIRQVRFABELKAIQEPDDARDYFQVAR 120
Db 61 VPLSDALLNKLGRYSQAIEALPPDKYQNESFAIRQVRFABELKAIQEPDDARDYFQVAR 120
QY 121 ANCKKFAFVHISFAOFELSCGNVKKSKOLLQKAVERGAVPLEMLBIALNRLNKKQLLS 180
Db 121 ANCKKFAFVHISFAOFELSCGNVKKSKOLLQKAVERGAVPLEMLBIALNRLNKKQLLS 180
QY 181 EBEKKNLSASTVLTQAESFSGSLGHLQNRNNSCDRGQTTKARFLYGENMPPQDAEIGYR 240
Db 181 EBEKKNLSASTVLTQAESFSGSLGHLQNRNNSCDRGQTTKARFLYGENMPPQDAEIGYR 240
QY 241 NSLRQTNKTKQSCPPGRVFNLLNSPDCDVKTTDSDVVPFCFMKQTSRSCRDLVVPGSKP 300
Db 241 NSLRQTNKTKQSCPPGRVFNLLNSPDCDVKTTDSDVVPFCFMKQTSRSCRDLVVPGSKP 300
QY 301 SGNDSCELENLKVQNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360
Db 301 SGNDSCELENLKVQNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360
QY 361 PEVPESNQWQAKRSECEINONPAASNNHQIPELARKVNTQEKHTTEQPVFVSVKOS 420
Db 361 PEVPESNQWQAKRSECEINONPAASNNHQIPELARKVNTQEKHTTEQPVFVSVKOS 420
QY 421 PPISTKWFDPKSICTKTPSSNTLDDYMSCFRTPVVKNDFPPACOLSTPYGPACFOQOQH 480
Db 421 PPISTKWFDPKSICTKTPSSNTLDDYMSCFRTPVVKNDFPPACOLSTPYGPACFOQOQH 480
QY 481 QILATPLQNLQVLASSANECISVKGRYISILKQIGSGSKVQVNLNKKQIYAIKYN 540
Db 481 QILATPLQNLQVLASSANECISVKGRYISILKQIGSGSKVQVNLNKKQIYAIKYN 540
QY 541 LEEADNQTLDYSYRNEIAYLNKLQOHSDKIIRLYDYETDQIYVMWECGNIDLSWLKXK 600
Db 541 LEEADNQTLDYSYRNEIAYLNKLQOHSDKIIRLYDYETDQIYVMWECGNIDLSWLKXK 600
QY 601 KSIDPWERKSYWKNMLEAVHTIHQGHVHSDLPANFLIVDGMKLIDFGIANQMPDPT 660
Db 601 KSIDPWERKSYWKNMLEAVHTIHQGHVHSDLPANFLIVDGMKLIDFGIANQMPDPT 660
QY 661 SVVKDSQVGTVMYMPPEAIKDMSSRENGSKSKISPKSDVWSLGCILYMTYKTPFQ 720
Db 661 SVVKDSQVGTVMYMPPEAIKDMSSRENGSKSKISPKSDVWSLGCILYMTYKTPFQ 720
QY 721 IINOISKLHAIIDPNHEIEFPDIPEKQDVLKCLKRPKQIRISIPELLAHPYVQIOTH 780
Db 721 IINOISKLHAIIDPNHEIEFPDIPEKQDVLKCLKRPKQIRISIPELLAHPYVQIOTH 780
QY 781 PVNOMAKGTTTEEMKYVGLQVGLNSPNSILKAAKTLYEHYSGGSHNSSSKTFEKKRGK 840
Db 781 PVNOMAKGTTTEEMKYVGLQVGLNSPNSILKAAKTLYEHYSGGSHNSSSKTFEKKRGK 840
QY 841 K 841
Db 841 K 841

RESULT 7
US-10-295-027-865
; Sequence 865, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 0:8501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 865
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-865

Query Match 100.0%; Score 4376; DB 15; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.4e-299;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	Db	Qy	Db	Qy	Db	Qy	Db	Qy	Db		
1	MNKVRDINKKFKNEDLDELNLNKISADTTDNGSTVNOIMMANNPEDMLSLLLKLEKNS	60	1	MNKVRDINKKFKNEDLDELNLNKISADTTDNGSTVNOIMMANNPEDMLSLLLKLEKNS	60	61	VPLSDALLNLKIGYSAIEALPPDKYQONESFARIOVRFAELKATCEPDDARDYFQMAR	120	61	VPLSDALLNLKIGYSAIEALPPDKYQONESFARIOVRFAELKATCEPDDARDYFQMAR	120
121	ANCKKFAFVHSFAQFELSGNKKSKQLLQKAVEGAVPLEMLETALRNLNKQKLLS	180	121	ANCKKFAFVHSFAQFELSGNKKSKQLLQKAVEGAVPLEMLETALRNLNKQKLLS	180	181	EEKKNLSASTVLTAAQSFSGSLGLHQRNNNSCDRSGQTTKARFLYXGNMPPQDAEIGYR	240	181	EEKKNLSASTVLTAAQSFSGSLGLHQRNNNSCDRSGQTTKARFLYXGNMPPQDAEIGYR	240
241	NSLRQTNKTKQSCFPGFVNNLNSPDCDKVKTDDSVVPCFMKQTSRSECRDLVVGSGYR	300	241	NSLRQTNKTKQSCFPGFVNNLNSPDCDKVKTDDSVVPCFMKQTSRSECRDLVVGSGYR	300						

RESULT 8

US-10-354-358-30
; Sequence 30, Application US/10354358
; Publication No. US20030157082A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Hunter, John Joseph
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Lesoon, Andrea
; APPLICANT: Lightcap, Eric S.
; APPLICANT: Williamson, Mark
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
; TITLE OF INVENTION: 7181, 7660, 25642, 69583, 49863, 8897, 1682, 17667, 9235,
; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 13427, 2160,
; TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
; TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
; TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
; FILE REFERENCE: MP102-020P1ENONNIM
; CURRENT APPLICATION NUMBER: US/10/354,358
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US 60/353,600
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/364,517
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/371,075
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/371,507

Db 77 VPLSDALLNKLIGRYSAIEALPPDKYQNESFARIQVRFAELKAIQEPDARDYFQMAR 136
QY 121 ANCKKFAFVHISFAQFELSGNKKVKSQKLOKAVERGAVPLEMLEIALRNLNLOKKKLS 180
Db 137 ANCKKFAFVHISFAQFELSGNKKVKSQKLOKAVERGAVPLEMLEIALRNLNLOKKKLS 196
QY 181 EEEKKNLSASTVLTAAQSFSGSLGHLQNRNNSCDRGOTTKARFLYGENNPPQDAEIGYR 240
Db 197 EEEKKNLSASTVLTAAQSFSGSLGHLQNRNNSCDRGOTTKARFLYGENNPPQDAEIGYR 256
QY 241 NSLRQTNKTQSCFPGRVPVNLNSPCDVKTDDSVVPCFMKQTSRSECRDLVWPGSKP 300
Db 257 NSLRQTNKTQSCFPGRVPVNLNSPCDVKTDDSVVPCFMKQTSRSECRDLVWPGSKP 316
QY 301 SGNDSCELRLNLSVQNSHFKPEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYOE 360
Db 317 SGNDSCELRLNLSVQNSHFKPEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYOE 376
QY 361 PEVPESQKQWAKRSEKSECINQNPAASSNHQVPELARKVNTQKHTTPEQPVFSVSKOS 420
Db 377 PEVPESQKQWAKRSEKSECINQNPAASSNHQVPELARKVNTQKHTTPEQPVFSVSKOS 436
QY 421 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFPPACQLSTPYGQACFQOOQH 480
Db 437 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFPPACQLSTPYGQACFQOOQH 496
QY 481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVLNEKKQIYAIKYVN 540
Db 497 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVLNEKKQIYAIKYVN 556
QY 541 LEEADNQTLSYRNEIAYLNKLOHSDKIIRLYDYEITDQIYIYMWECGNIDLSWLKXK 600
Db 557 LEEADNQTLSYRNEIAYLNKLOHSDKIIRLYDYEITDQIYIYMWECGNIDLSWLKXK 616
QY 601 KSIDPWERKSWKMLKLEAVHTIHQGHVHSDLPANFLIVDGMKLKIDFGIANQMPDIT 660
Db 617 KSIDPWERKSWKMLKLEAVHTIHQGHVHSDLPANFLIVDGMKLKIDFGIANQMPDIT 676
QY 661 SVVKDSQVGTVMYPPPEAIKDMSSRENGSKSKISPKSDVWSLGCILYMTYKTPFQO 720
Db 677 SVVKDSQVGTVMYPPPEAIKDMSSRENGSKSKISPKSDVWSLGCILYMTYKTPFQO 736
QY 721 IINOISKLHAIIDPNHEIEPDIPEKLDQVLKCLKRDPKORISIPPELLAHYVQIOTH 780
Db 737 IINOISKLHAIIDPNHEIEPDIPEKLDQVLKCLKRDPKORISIPPELLAHYVQIOTH 796
QY 781 PVNQAKGTTTEEMKYVLQVLGNSPNSILKAAKTLYEHYSGESHSSSSKTFEKKRGK 840
Db 797 PVNQAKGTTTEEMKYVLQVLGNSPNSILKAAKTLYEHYSGESHSSSSKTFEKKRGK 856
QY 841 K 841
Db 857 K 857

RESULT 10
US-10-173-999-34
; Sequence 34, Application US/10173999
; Publication NO. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; TITLE OF INVENTION: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; PRIOR FILING DATE: 2002-06-17
; PRIOR FILING DATE: 2002-06-17
; PRIOR FILING DATE: 2001-06-18
; PRIOR FILING DATE: 2001-06-18
; PRIOR FILING DATE: 2001-08-27

; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: Patent In Ver: 2.1
; SEQ ID NO 34
; LENGTH: 857
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-173-999-34

Query Match 99.9%; Score 4373; DB 15; Length 857;
Best Local Similarity 99.9%; Pred. No. 4e-299;
Matches 840; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKYRDTKNKFNEDLTDELNLNKISADTTDNGTGVNQIMMANNPPEDWL-SLLLLKLEKNS 60
Db 17 MNKYRDTKNKFNEDLTDELNLNKISADTTDNGTGVNQIMMANNPPEDWL-SLLLLKLEKNS 76
QY 61 VPLSDALLNKLIGRYSAIEALPPDKYQNESFARIQVRFAELKAIQEPDARDYFQMAR 120
Db 77 VPLSDALLNKLIGRYSAIEALPPDKYQNESFARIQVRFAELKAIQEPDARDYFQMAR 136
QY 121 ANCKKFAFVHISFAQFELSGNKKVKSQKLOKAVERGAVPLEMLEIALRNLNLOKKKLS 180
Db 137 ANCKKFAFVHISFAQFELSGNKKVKSQKLOKAVERGAVPLEMLEIALRNLNLOKKKLS 196
QY 181 EEEKKNLSASTVLTAAQSFSGSLGHLQNRNNSCDRGOTTKARFLYGENNPPQDAEIGYR 240
Db 197 EEEKKNLSASTVLTAAQSFSGSLGHLQNRNNSCDRGOTTKARFLYGENNPPQDAEIGYR 256
QY 241 NSLRQTNKTQSCFPGRVPVNLNSPCDVKTDDSVVPCFMKQTSRSECRDLVWPGSKP 300
Db 257 NSLRQTNKTQSCFPGRVPVNLNSPCDVKTDDSVVPCFMKQTSRSECRDLVWPGSKP 316
QY 301 SGNDSCELRLNLSVQNSHFKPEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYOE 360
Db 317 SGNDSCELRLNLSVQNSHFKPEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYOE 376
QY 361 PEVPESQKQWAKRSEKSECINQNPAASSNHQVPELARKVNTQKHTTPEQPVFSVSKOS 420
Db 377 PEVPESQKQWAKRSEKSECINQNPAASSNHQVPELARKVNTQKHTTPEQPVFSVSKOS 436
QY 421 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFPPACQLSTPYGQACFQOOQH 480
Db 437 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFPPACQLSTPYGQACFQOOQH 496
QY 481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVLNEKKQIYAIKYVN 540
Db 497 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVLNEKKQIYAIKYVN 556
QY 541 LEEADNQTLSYRNEIAYLNKLOHSDKIIRLYDYEITDQIYIYMWECGNIDLSWLKXK 600
Db 557 LEEADNQTLSYRNEIAYLNKLOHSDKIIRLYDYEITDQIYIYMWECGNIDLSWLKXK 616
QY 601 KSIDPWERKSWKMLKLEAVHTIHQGHVHSDLPANFLIVDGMKLKIDFGIANQMPDIT 660
Db 617 KSIDPWERKSWKMLKLEAVHTIHQGHVHSDLPANFLIVDGMKLKIDFGIANQMPDIT 676
QY 661 SVVKDSQVGTVMYPPPEAIKDMSSRENGSKSKISPKSDVWSLGCILYMTYKTPFQO 720
Db 677 SVVKDSQVGTVMYPPPEAIKDMSSRENGSKSKISPKSDVWSLGCILYMTYKTPFQO 736
QY 721 IINOISKLHAIIDPNHEIEPDIPEKLDQVLKCLKRDPKORISIPPELLAHYVQIOTH 780
Db 737 IINOISKLHAIIDPNHEIEPDIPEKLDQVLKCLKRDPKORISIPPELLAHYVQIOTH 796
QY 781 PVNQAKGTTTEEMKYVLQVLGNSPNSILKAAKTLYEHYSGESHSSSSKTFEKKRGK 840
Db 797 PVNQAKGTTTEEMKYVLQVLGNSPNSILKAAKTLYEHYSGESHSSSSKTFEKKRGK 856
QY 841 K 841

Db 857 K 857

RESULT 11

US-10-664-421-157

Sequence 157, Application US/10664421

Publication No. US20040142864A1

GENERAL INFORMATION:

APPLICANT: BREWER, RYAN

APPLICANT: IBRAHIM, PRAHA

APPLICANT: KUNAR, ABHINAV

APPLICANT: MANDIYAN, VALSAN

APPLICANT: MILBURN, MICHAEL V.

TITLE OF INVENTION: CRYSTAL STRUCTURE OF PIM-1 KINASE

FILE REFERENCE: 039363/0703

CURRENT APPLICATION NUMBER: US/10/664,421

CURRENT FILING DATE: 2003-09-16

PRIOR APPLICATION NUMBER: 60/412,341

PRIOR FILING DATE: 2002-09-20

PRIOR APPLICATION NUMBER: 60/411,398

PRIOR FILING DATE: 2002-09-16

NUMBER OF SEQ ID NOS: 169

SOFTWARE: PatentIn Ver. 3.2

SEQ ID NO 157

LENGTH: 367

TYPE: PRT

ORGANISM: Homo sapiens

US-10-664-421-157

Query Match 44.1%; Score 1929; DB 16; Length 367;

Best Local Similarity 100.0%; Pred. No. 1.7e-127; Indels 0;

Matches 367; Conservative 0; Mismatches 0; Gaps 0;

QY 459 FPPACQLSTPYGQACFQQQHQHILATPLQNLQVLASSANECISVKGRIYSILKQIGSG 518

Db 1 FPPACQLSTPYGQACFQQQHQHILATPLQNLQVLASSANECISVKGRIYSILKQIGSG 60

QY 519 GSSKVFQVLNEKKIYAIKYVNLBEADNQLDSYRNEIAYNLKQHSKIIILYDEIT 578

Db 61 GSSKVFQVLNEKKIYAIKYVNLBEADNQLDSYRNEIAYNLKQHSKIIILYDEIT 120

QY 579 DQYIYVWMECGNIDLSNWLKKGKSIDPWERKSYWKNLEAVHTIHOIGIVHSDLKPFANFL 638

Db 121 DQYIYVWMECGNIDLSNWLKKGKSIDPWERKSYWKNLEAVHTIHOIGIVHSDLKPFANFL 180

QY 639 IVDGMLKLIDFGIANQMPDTSVVKDSQVGTVMYMPPEAIKDMSSSRENGKSKISPK 698

Db 181 IVDGMLKLIDFGIANQMPDTSVVKDSQVGTVMYMPPEAIKDMSSSRENGKSKISPK 240

QY 699 SDVWSLGCILYMTYGTPTFQIINOISKLHAIIDPNHEIEFPDIPKDIQDVLKCKLKR 758

Db 241 SDVWSLGCILYMTYGTPTFQIINOISKLHAIIDPNHEIEFPDIPKDIQDVLKCKLKR 300

QY 759 DPQRISIPPELLAHYPYQIOTHPVNMAGKGTTEEMKYVLGQVLGNSPNSILKAAKTYE 818

Db 301 DPQRISIPPELLAHYPYQIOTHPVNMAGKGTTEEMKYVLGQVLGNSPNSILKAAKTYE 360

QY 819 HYSGES 825

Db 361 HYSGES 367

RESULT 12

US-10-081-119-32

Sequence 32, Application US/10081119

Publication No. US20030045491A1

GENERAL INFORMATION:

APPLICANT: Reinhard, Christoph

APPLICANT: Jefferson, Anne B.

APPLICANT: Chan, Vivien W.

TITLE OF INVENTION: TTK in diagnosis and as a Therapeutic

TITLE OF INVENTION: Target in Cancer

FILE REFERENCE: 16932.002

CURRENT APPLICATION NUMBER: US/10/081,119

CURRENT FILING DATE: 2002-02-21

PRIOR APPLICATION NUMBER: 60/289,813

PRIOR FILING DATE: 2001-02-21

NUMBER OF SEQ ID NOS: 38

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 32

LENGTH: 678

TYPE: PRT

ORGANISM: Schizosaccharomyces pombe

US-10-081-119-32

Query Match 15.4%; Score 673.5; DB 14; Length 678;

Best Local Similarity 30.9%; Pred. No. 1.1e-38;

Matches 195; Conservative 102; Mismatches 210; Indels 125; Gaps 23;

QY 233 QDAEICYRNSLPQTNTKQSCFPGRVNVNLLNSPCDVKTDSVVPFCMKQKTSRSECD 292

Db 32 QDPELYFKN---DTFSKSS-----HSDGTVTGDTLRKQSSGATALE 70

QY 293 LVV--PGSK-----PSGNDSCELR--NLKSVQNSHFKEPLVSDKSSSELIITDSITLKNK 343

Db 71 RLVSHPRTKNFDLQNGGQNSALKEVNTPAYGSMHFEHLITPLPS-----TNASH 121

QY 344 TESSLLAKLEETKEYQEPB-VPESNQK-----QWAKKSECCINONPAASNNHWOIPELA 397

Db 122 SEVLSAGVNDLNSSEHDLPLKSVNKTFGSLISRRRRIRIGLGP-----PKRA 172

QY 398 R-----KVNTQKHHTTPEQPVFSVSKQSPPISTF---KWFDPKSICTPSSNTLDDVM 447

Db 173 EYTLTDPKSTDKNSTADEDIEMKSRVSPASNSVAATLTKPLQLHNTPLQTSQEHPK 232

QY 448 SCFRTPVKNDFPPACQL-----STPYGQP-ACFOQQQ-----HOI----- 482

Db 233 PSFHPSQFESSFSPRVQFDHVERRASELHSPVTVFQEPQSRASQPYESHALSPPKAPL 292

QY 483 ----LATPLQNLQVLASSANECISVKGRIYSILKQIGSGSKVFOVLN-EKKOIYATK 537

Db 293 FNSQATPIPKQ-----QDVIVANLQFKLVGKGGSSWYRIFSPNSRLYALK 345

QY 538 YVNLBEADNQLDSYRNEIAYNLKQHSKIIILYDEITDQ--YIYVWMECGNIDLSN 595

Db 346 EYVNFADQTTIOGKVEIALLRKL-SGNDRIIKLYAAEVNDTLQGLNVMCEGTDLAN 404

QY 596 WLKK--KKSIDPWERKSYWKNLEAVHTIHOIGIVHSDLKPFANFLVDGMLKLIDFGIAN 653

Db 405 LLMKNMKPFINLNFIRMYWEQMLEAVQVVDQNIHVHSDLKPFANFLVEGNLKLIDFGIAK 464

QY 654 QMOPDTSVVKDSQVGTVMYMPPEAIKDMSSSRENGKSKISPKSDVWSLGCILYMTY 713

Db 465 AIGDNTTNIHRDISHGTINYVAPEALTDMNAHTNSGVKLKGRPSDVWSLGCILYQVY 524

QY 714 GKTPTQIINOISKLHAIIDPNHEIEFPD-----IPEK-----DLQVVKCC 755

Db 525 GRAPEAH-LKMIQAIAPNEQYHIFPEVALPANAVQEKESGLPGVTGPDLMVDMKRC 583

QY 756 LKRDPKQRISIPPELLAHYPYQIOTHPVNMAGK 787

Db 584 LERDQKRLLIIFELLVHFPFLNPLPSYLPLAK 615

RESULT 13

US-10-437-963-201272

Sequence 201272, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.


```
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 43102
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700729819_FLI.pep
US-10-425-114-43102
```

```
Query Match      14.0%; Score 614; DB 12; Length 399;
Best Local Similarity 41.1%; Pred. No. 8.1e-35;
Matches 139; Conservative 53; Mismatches 106; Indels 40; Gaps 6;

Qy      504 VKGRIYSILKQIGSGGSKYFOVLNEKKQIYAIKYNLEADNQTLDSYENEIAYLNKLO 563
Db      16  VNGKLYQRGLGIGSGSEVHKVVISDCTIYALKRIKGRDYATAYGFCQEIEYLNRL- 74

Qy      564 QHSDKIIRLYDYETDQ-----YIYVMECGNIDLNSWLKKK- 600
Db      75  KGKNNIIQLIDYEYTDKALPEGVINGSFNKDGKRVKDDGYIYVLEVEYGEIDLAHMLSQKW 134

Qy      601 -----KSIDPWERKSYWKNMLEAVYTHQHGIHVSDDLXPANFLIVDGMKLLIDFGIAN 653
Db      135 KELDGSNQITIDENWLRFTYWOQILOAVNTIHEERIVHSDLXPANFLLVKGSKLIDFGIAK 194

Qy      654 QMQPDTT-SVVKDSQVGTWVYMPPEAIKDMSSSRRENGSKSKISPKSDVWSLGCILYYWY 713
Db      195 AIMSDDTNIQEDSQVGTLSYMSPEAF--MCNESDASGNIKCGRPDSINSLGCILYQWY 252

Qy      714 GKTPFQQLIINQISKLHAIIDPNHIEIPDPEKDLQDLVKCCCLKRDPKQKQISIPPELLAHP 773
Db      253 GRTPFADYKTFWAKFKVITDPNHEIMYAPVSNPWLLDMRCLAWDRNERWRIPQLIQHP 312

Qy      774 YVQICTHPVQWAKGTTTEEMKYVLGQLVG---LNSPNS 808
Db      313 FLV-----PPVPCHSFSFYQEQTKFLQLLISEACINDPES 346
```

Search completed: September 29, 2004, 16:30:17
Job time : 148 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 16:03:10 ; Search time 27 Seconds
(without alignments)
1621.889 Million cell updates/sec

Title: US-10-081-119-14
Perfect score: 4376
Sequence: 1 MNKVRDIKNKFKQEDLTDEL.....GGESHNSSSSKTFKKRGK 841

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4376	100.0	841	1 TTK_HUMAN	P33981 homo sapien
2	3165.5	72.3	856	1 TTK_MOUSE	P35761 mus musculus
3	673.5	15.4	678	1 MPH1_SCHPO	O34235 schizosacch
4	586.5	13.4	764	1 MFS1_YEAST	P54199 saccharomyc
5	347.5	7.9	772	1 TLK2_HUMAN	Q86488 homo sapien
6	335	7.7	718	1 TLK2_MOUSE	O55047 mus musculus
7	332	7.6	766	1 TLK1_HUMAN	Q86488 homo sapien
8	328	7.5	766	1 TLK1_MOUSE	Q86488 homo sapien
9	327	7.5	460	1 CSN1_SCHPO	O39170 schizosacch
10	324.5	7.4	576	1 SNF1_SCHPO	O74536 schizosacch
11	318.5	7.3	1607	1 MK4_HUMAN	Q97614 homo sapien
12	309	7.1	1231	1 YKT3_CAEEL	P34314 caenorhabdi
13	304.5	7.0	1401	1 WIS4_SCHPO	O14299 schizosacch
14	303.5	6.9	1597	1 M3K4_MOUSE	O08648 mus musculus
15	299	6.8	1478	1 BCK1_YEAST	Q01389 saccharomyc
16	295	6.7	1080	1 NPK1_YEAST	P38692 saccharomyc
17	294.5	6.7	515	1 STE7_YEAST	P06784 saccharomyc
18	294.5	6.7	631	1 KOKO_MOUSE	Q60806 mus musculus
19	294.5	6.7	1081	1 KOKO_YEAST	Q12236 saccharomyc
20	294.5	6.7	1142	1 PKA1_YEAST	P38990 saccharomyc
21	292	6.7	1914	1 KMLS_HUMAN	O15746 homo sapien
22	289.5	6.6	544	1 PAK3_MOUSE	Q61036 mus musculus
23	288.5	6.6	610	1 CDPL1_ARATH	O06850 arabidopsi
24	288.5	6.6	620	1 SNF1_CANAL	P52497 candida alb
25	287.5	6.6	730	1 CDK9_CAEEL	P46551 caenorhabdi
26	287	6.6	988	1 PKL1_SCHPO	P36582 schizosacch
27	285.5	6.5	545	1 PKA1_MOUSE	O08643 mus musculus
28	285.5	6.5	1906	1 KMLS_CHICK	P11799 gallus gall
29	285	6.5	615	1 CNK_RAT	Q97011 rattus norv
30	282.5	6.5	443	1 ST24_HUMAN	Q97600 homo sapien
31	282.5	6.5	544	1 PAK3_HUMAN	O75914 homo sapien
32	282.5	6.5	646	1 CNK_HUMAN	Q944b4 homo sapien
33	281.5	6.4	491	1 STK3_HUMAN	Q31388 homo sapien

ALIGNMENTS

RESULT 1

TTK_HUMAN

ID	TTK_HUMAN	STANDARD;	PRT;	841 AA.
AC	P33981;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Dual specificity protein kinase TTK (EC 2.7.1.1-) (PVT).			
GN	TTK OR MPS11.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-92348472; PubMed-1639825;			
RA	Mills G.B., Schmandt R., McGill M., Amendola A., Hill M.,			
RA	Jacobs K., May C., Rodricks A.-M., Campbell S., Hogg D.,			
RT	"Expression of TTK, a novel human protein kinase, is associated with			
RT	cell proliferation.";			
RL	J. Biol. Chem. 267:16000-16006(1992).			
RN	[2]			
RP	SEQUENCE OF 509-776 FROM N.A.			
RX	MEDLINE-92065863; PubMed-1956325;			
RA	Lindberg R.A., Hunter T.;			
RL	Unpublished results, cited by:			
RL	Hanks S.K., Quinn A.M.;			
RL	Meth. Enzymol. 200:38-62(1991).			
CC	-I- FUNCTION: PHOSPHORYLATES PROTEINS ON SERINE, THREONINE, AND			
CC	TYROSINE. PROBABLY ASSOCIATED WITH CELL PROLIFERATION.			
CC	-I- TISSUE SPECIFICITY: PRESENT IN RAPIDLY PROLIFERATING CELL LINES.			
CC	-I- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; M86699; AAA61239.1; -			
DR	PIR; A42861; A42861.			
DR	HSSP; Q00534; 1B18.			
DR	Genew; HGNC:12401; TTK.			
DR	NTM; 604092; -			
DR	GO; GO:0005819; C.spindle; TAS.			
DR	GO; GO:0004674; F.protein serine/threonine kinase activity; TAS.			
DR	GO; GO:0004713; F.protein-tyrosine kinase activity; TAS.			
DR	GO; GO:0007052; P.mitotic spindle assembly; TAS.			
DR	GO; GO:0007094; P.mitotic spindle checkpoint; TAS.			
DR	GO; GO:0008284; P.positive regulation of cell proliferation; TAS.			
DR	GO; GO:0000074; P.regulation of cell cycle; TAS.			
DR	InterPro; IPR000719; Prot_kinase.			
DR	InterPro; IPR008271; Ser_thr_pkin_AS.			
DR	InterPro; IPR002290; Ser_thr_kinase.			

34	280.5	6.4	544	1 PAK3_RAT	Q62829 rattus norv
35	280	6.4	891	1 KIN1_SCHPO	P22987 schizosacch
36	279.5	6.4	544	1 PAK1_RAT	P35465 rattus norv
37	279	6.4	387	1 KPSG_MOUSE	P07934 mus musculus
38	278	6.4	387	1 KPSG_RAT	P13286 rattus norv
39	278	6.4	586	1 KMS8_YEAST	Q03533 saccharomyc
40	277.5	6.3	893	1 KN9R_YEAST	P53739 saccharomyc
41	277	6.3	313	1 NEK5_HUMAN	Q9hc98 homo sapien
42	277	6.3	313	1 NEK6_MOUSE	Q9es70 mus musculus
43	277	6.3	408	1 STGL_XENLA	Q91819 xenopus lae
44	276.5	6.3	384	1 ARK1_SCHPO	O59790 schizosacch
45	276.5	6.3	511	1 NEK3_MOUSE	Q9roa5 mus musculus

```
DR InterPro: IPR008941; TPR-like.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TK; 1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
DR PROSITE: PS00111; PROTEIN KINASE DOM; 1.
DR Transferrase: Serine/threonine-protein kinase; ATP-binding;
KW Tyrosine-protein kinase.
FT DOMAIN 509 775 PROTEIN KINASE.
FT NP_BIND 515 523 ATP (BY SIMILARITY).
FT BINDING 537 537 ATP (BY SIMILARITY).
FT ACT_SITE 631 631 BY SIMILARITY.
FT CONFLICT 752 752 L -> V (IN REF. 2).
SQ SEQUENCE 841 AA; 95279 MW; ADEEC2ACD15EF08C CRC64;

Query Match 100.0%; Score 4376; DB 1; Length 841;
Best Local Similarity 100.0%; Pred. No. 1.4e-213;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKYRDIKNKFKNEDLTDELSLNKISADTTDNGSTVQINMMANNPEDWLSLLKLEKNS 60
DB 1 MNKYRDIKNKFKNEDLTDELSLNKISADTTDNGSTVQINMMANNPEDWLSLLKLEKNS 60

QY 61 VPLSDALLNKILGRYSQAIEALPPDKYQONESFARIQVRFPAELKAIQEPDADYFQMAR 120
DB 61 VPLSDALLNKILGRYSQAIEALPPDKYQONESFARIQVRFPAELKAIQEPDADYFQMAR 120

QY 121 ANCKKFAFVHSFAQFELSQQNVKSKQLLOKAVERGAVLEMLELALRNINLQKQLLS 180
DB 121 ANCKKFAFVHSFAQFELSQQNVKSKQLLOKAVERGAVLEMLELALRNINLQKQLLS 180

QY 181 BEEKKNLSASVLTAAQSFSGSLGHLQNRNNSCDRGQTTKARFLYGENNPPDABIGYR 240
DB 181 BEEKKNLSASVLTAAQSFSGSLGHLQNRNNSCDRGQTTKARFLYGENNPPDABIGYR 240

QY 241 NSLRQTNKTKQSCFGRVFNVLNNSPDCVKTDDSVVPCMKQTSRSECRDVLVPGSKP 300
DB 241 NSLRQTNKTKQSCFGRVFNVLNNSPDCVKTDDSVVPCMKQTSRSECRDVLVPGSKP 300

QY 301 SGNDSCELRNLSQNSHFKEPLVSDKSSSELIITDSITLKNKTESLAKLEETKEYOE 360
DB 301 SGNDSCELRNLSQNSHFKEPLVSDKSSSELIITDSITLKNKTESLAKLEETKEYOE 360

QY 361 PEVPESKQKQAKRKSECINQNPASSNHQIPELARKVNTBQKHTTFEPQPVFSVKQS 420
DB 361 PEVPESKQKQAKRKSECINQNPASSNHQIPELARKVNTBQKHTTFEPQPVFSVKQS 420

QY 421 PPISTSKWFDPKSICKTPSSNTLDDYMSCRTPVKNDRPPACOLSTPYGOPACFOQOQH 480
DB 421 PPISTSKWFDPKSICKTPSSNTLDDYMSCRTPVKNDRPPACOLSTPYGOPACFOQOQH 480

QY 481 QILATPLQNQLQVLASSANECISVKGRISYILKQIGSGGSSKVFQVLNEKKQIYAIKYVN 540
DB 481 QILATPLQNQLQVLASSANECISVKGRISYILKQIGSGGSSKVFQVLNEKKQIYAIKYVN 540

QY 541 LEADNQTLDYSRNEIAYLNKLOHSDKIIRLDYDYEITDQYIYVMECCGNIDLSNWLKX 600
DB 541 LEADNQTLDYSRNEIAYLNKLOHSDKIIRLDYDYEITDQYIYVMECCGNIDLSNWLKX 600

QY 601 KSIDPWERKSWYKMKLEAVHTIHGHGVHSDLPANFLIVDGMKLIIDFGIANQMPDIT 660
DB 601 KSIDPWERKSWYKMKLEAVHTIHGHGVHSDLPANFLIVDGMKLIIDFGIANQMPDIT 660

QY 661 SVVKDSQVGTVMYPPPAIKDMSSRENGSKSIKSPKSDVMSGLGILYMTYTKTPFQ 720
DB 661 SVVKDSQVGTVMYPPPAIKDMSSRENGSKSIKSPKSDVMSGLGILYMTYTKTPFQ 720

QY 721 IINQISKLHAIIDNHEIEPDPIDKLDQVLKCLKRDPKQIRISIPELLAHPYVQIQH 780
DB 721 IINQISKLHAIIDNHEIEPDPIDKLDQVLKCLKRDPKQIRISIPELLAHPYVQIQH 780

QY 781 PVNQMAKGTTEMKYVJGQLVGLNPSNLSILKAATLYEHYSGGSHSSSKTTEKRGK 840
```

```
DB 781 PVNQMAKGTTEMKYVJGQLVGLNPSNLSILKAATLYEHYSGGSHSSSKTTEKRGK 840
QY 841 K 841
DB 841 K 841

RESULT 2
TTK MOUSE
ID TTK MOUSE STANDARD; PRT; 856 AA.
AC P35761;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dual specificity protein kinase TTK (EC 2.7.1.-) (ESK) (PYT).
GN TTK OR ESK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=92269841; PubMed=1375325;
RA Douville E.M.J., Afar D.E.H., Howell B.W., Letwin K., Tannock L.,
RA Ben-David Y., Pawson T., Bell J.C.;
RT "Multiple cDNAs encoding the esk kinase predict transmembrane and
RT intracellular enzyme isoforms."
RL Mol. Cell. Biol. 12:2681-2689(1992).
CC -!- FUNCTION: PHOSPHORYLATES PROTEINS ON SERINE, THREONINE, AND
CC TYROSINE. MAY PLAY SOME ROLE IN THE CONTROL OF CELL PROLIFERATION
CC OR DIFFERENTIATION AND COULD BE INVOLVED IN MODULATING DIFFERENT
CC LEVELS OF SIGNAL TRANSDUCTION PATHWAYS.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P35761-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P35761-2; Sequence=VSP_004875;
CC -!- TISSUE SPECIFICITY: PRESENT IN RAPIDLY PROLIFERATING CELL LINES;
CC HIGH LEVELS IN TESTIS, BONE MARROW, SPLEEN AND THYMUS. LOW LEVELS
CC IN BRAIN, HEART, LUNG AND KIDNEY.
CC -!- PTM: Autophosphorylated.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC EMBL; M86377; AAA37578.1; -.
DR PIR; A44439; A44439.
DR PIR; B44439; B44439.
DR HSP; Q00334; IBI8.
DR MGI; MGI:119421; Ttk.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR008941; TPR-like.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Tyrosine-protein kinase; Alternative splicing; Phosphorylation.
FT DOMAIN 524 790 PROTEIN KINASE.
FT NP_BIND 530 538 ATP (BY SIMILARITY).
FT BINDING 552 552 ATP (BY SIMILARITY).
```

```
FT ACT SITE 546 546 BY SIMILARITY.
FT VARSPLIC 296 321 Missing (in isoform 2).
FT /FTID=VSP_004875.
SQ SEQUENCE 856 AA; 96210 MW; 31f8c16c195e5e96 CRC64;

Query Match
Best Local Similarity 71.8%; Score 3165.5; DB 1; Length 856;
Matches 622; Conservative 86; Mismatches 107; Indels 51; Gaps 8;

Qy 1 MNKVRDINKFKNEEDLDELISLNKISADTTDNGTQVNIQIMMANNPDLWSLLKLEKKS 60
Db 17 MSKVRDINKKI-NEDCTDELSLKICADHTE---TVNQIMRVGNTPEPNLNFLLKLEKKS 72
Qy 61 VPLSDALLNLKILGRYSOAIKALPDYKQNESFARIOVFAELKATICEPDADYFQMAR 120
Db 73 SPLNDDLLNLKILGRYSOAIKALPDYKQNESFARIOVFAELKATICEPDADYFQMAR 132
Qy 121 ANCKKFAFVHISFAQFELSOGNVKSKQLLOKAVERGAVPLEMLEIALRNLNLKQKQLLS 180
Db 133 ENCKKFAFVHISFAQFELSOGNVKSKQLLOKAVERGAVPLEMLEIALRNLNLKQKQLLP 192
Qy 181 EEEKKNLSASTVLTQSFSGSLGHONRNNSCDSRGOTTAKPFLYGENPPODAELGYR 240
Db 193 EEDKXSVASATVLSAQPFSSSLGNVQNRISCSERQAGAAVLYGENLPPODAEVRHQ 252
Qy 241 NSLRQTNKTQSCFGRVPVNLNSPDCDKVTDDSV-----VPCFMKR 283
Db 253 NPFQKTHAKGSCFGRVPVNLNSPDPFYKTDSSAVTQLTTRLALSSVPLPYVTCLLHL 312
Qy 284 Q-----TSRSRCDRLVPGSPGNSDCELRNLKSVQNSHPKEPLVDEKSSSELIIT 335
Db 313 QLLALAGLAGSGPDRDAILPGSRPGSDSYELRGLKPIQITVYKDSLVNEKSSSEL-MS 371
Qy 336 DSITLKNKTESLAKLEETKEYCEPVPESNQOWAKKSEKINONPAASNNHWOIPE 395
Db 372 DLIALKSKTDS-IJKLEETK----PEIARPNQWSTKPKPCVFQNPAPFLRHVPD 426
Qy 396 LARKVNTQKHTTEQPVFSVSKSPPISTSKWDPKSIKCTPSSNTLDYMSCFPTPVV 455
Db 427 VTPRAD-----KSPPTISVPKMLDPKSACTETPSSSLDDYMKCFKTPPV 470
Qy 456 KNDPPACOLSTPYGPACQOQOQOHLATPLQNLQVLASSANECSVKGRIYSILKQI 515
Db 471 KNDPPACPSSTPYSQLARQOQOQOGLSTPLQSLQSGSSINECSVNGRIYSILKQI 530
Qy 516 GSGGSSKVFQVLEKKQIYAIKYVNLEADNQTLDYSRNEIAYLNKLQOHSDKIIRLYDY 575
Db 531 GSGGSSKVFQVLEKKQIYAIKYVNLEADNQTLDYSRNEIAYLNKLQOHSDKIIRLYDY 590
Qy 576 BITQYIYMWECGNIDLSWLKXKSIDPWERKSYWNKMLEAVHTHQHGIHVSIDLKPA 635
Db 591 EITEQYIYMWECGNIDLSWLKXKSIDPWERKSYWNKMLEAVHTHQHGIHVSIDLKPA 650
Qy 636 NFLIVDGMKLIIDFGIANOMPDTSVVKDSQVGTVMYMPPEAIKDMSSSRENGKSKSI 695
Db 651 NFVIVDGMKLIIDFGIANOMPDTSIVKDSQVGTVMYMPPEAIKDMSSSRENGKSKSI 710
Qy 696 SPKSDVWSLGCILYMYTGTPTFOQINQISKLHALIDPNHEIFBFDIPKDLQDVLUKCC 755
Db 711 SPKSDVWSLGCILYMYTGTPTFOQIHINQVSKLHALINPAHEIFBFDIPKDLQDVLUKCC 770
Qy 756 LKRPKQKISIPELLAHPEYVIOIOTHPVNMOKAGTTEENKYVLGOLVGNPSNLTAKKT 815
Db 771 LVNPKRERISIPELLTHPEYVIOIQHPGSGMARGATDENKYVLGOLVGNPSNLTAKKT 830
Qy 816 LYENYSGGESHSNSSSKTFFKKRQKK 841
Db 831 LYERNCGEGGSSSKTFFDKRERK 856

RESULT 3
MPH1_SCHPO
ID MPH1_SCHPO STANDARD; PRT; 678 AA.
```

```
AC O94235; Q9P7Z5; Q9URW0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine/threonine protein kinase mph1 (EC 2.7.1.37).
OS MPH1 OR SPBC106.01 OR SPBC243.01 OR SPBC1271.16C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
SEQUENCE FROM N.A., AND FUNCTION.
RP MEDLINE=98264560; PubMed=9601094;
RA He X., Jones M.H., Winey M., Sazer S.;
RT "Mph1, a member of the Mps1-like family of dual specificity protein
RT kinases, is required for the spindle checkpoint in S. pombe.";
RL J. Cell Sci. 111:1635-1647(1998).
RN [2]
SEQUENCE FROM N.A.
RP STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feitwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weljens J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Cabel C., Fuchs M., Fricz C., Holzer E., Voest D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleau V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: Involved in the regulation of the onset of mitosis.
CC Involved in a pathway that coordinates cell proliferation and
CC differentiation.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AF020705; RAD01648.1; -
CC EMBL; AL133362; CAB62431.1; -
CC EMBL; AL034353; CAB72266.1; -
CC EMBL; AL110295; CAB53717.1; -
CC PIR; T43539; T43539.
CC GeneDB Spombe; SPBC106.01; -
CC GO; GO:0016043; P:cell organization and biogenesis; ISS.
CC GO; GO:0000074; P:regulation of cell cycle; ISS.
CC GO; GO:0007165; P:signal transduction; ISS.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_thr_pkin_AS.
```



```
FT ACT SITE 563 563 BY SIMILARITY.
FT MUTAGEN 580 580 D-SA: LOSS OF ACTIVITY.
FT CONFLICT 146 146 A -> S (IN REF. 1).
FT CONFLICT 211 213 RRE -> TKR (IN REF. 1).
SQ SEQUENCE 764 AA; 86827 MW; 26B171DFB88EB4D CRC64;

Query Match 13.48; Score 586.5; DB 1; Length 764;
Best Local Similarity 26.11; Pred. No. 1.3e-22; Indels 161; Gaps 27;
Matches 188; Conservative 140; Mismatches 230;

QY 178 LLSSEKKNLASASTVLTAQESFS-----GSLGHLQNANNSCDRGQT 219
DQ 40 LLSHTEK-TSASBILSHNNDKTANLEEMDRSSSRHPPSPMGNTLSGHTSTSH-ST 96
QY 220 TKARFLVGN-----NPPQDAEYCYNS-----LRQ----- 245
DQ 97 LFGYLRLNNHOTSMWTNTSDIEINVGNLSDKSPERINLRONKWKEDITAKYAEKSRKF 156
QY 246 --TNKTKQSCPFGRVPV-----NLLNSPDCDVKTDDSV-VPCFMKQRTSRSECDLVV 295
DQ 157 LISNRTTKLGPAKRAMTLTNIFDEVDVFNPNQINARETVLEPLEDSDHQINFKERRE-- 213
QY 296 PGSKPSGNDSCELNLSVQNSHPKE-----PLVSD-----EKSS 330
DQ 214 ---NTDYSDIDFGDLNPIQ--YIKGNLPTSDLPILISQIYFDKQREENQAALRHSSR 267
QY 331 ELIITDSTLTKNKTESSLLAKLEETKEVQBEVPEVSNQKQWQAKKQCEKINQPAASNNH 390
DQ 268 ELLYKSRSSSSLSNNLLAN-----KNSITSNNGS--QPRK---VSTGSSSKSS 315
QY 391 WQIPELARKVNTQKHTTFEQPVFSVSKQSPITSTKWF--PKSICKTSSN----- 441
DQ 316 IEIRALKENIDTGNNSNFNSPIHKYK--GISRNKDSSEKEVLRNISINAHADNL 372
QY 442 -----TLDDYMSCFRPVVKVNDPPACQILSTPYGQACFQQQOHOI-LATPLQNL 490
DQ 373 LQENKRLKRLDD-----ALTNNENINSKNLEVFYHRPAPPPVTKVEIVPEPKSA 424
QY 491 QVLASSANECISVKGRYISLKQIGSGSKKVPQVLNKKQIYAIKYVNALEADNQTLD 550
DQ 425 SL---SNRNITVNDSDYKIELLGRGSSRVYKVGSGNRVYALKRVDFDADSSID 481
QY 551 SYRNEIAYLNKQOHSDKIILYDIETDQVIYMWMEGNIDLSWLKKKS--IDPWER 608
DQ 482 GFKGEIELLEKLDQK-RVIQLDYEMGDGLYIMECGDHLSCQILNQRSGMPLDNFV 540
QY 609 KSYKNWMLAEVHTTHQGVHSDLKPANFLVDGMLKLIIDFGIANQWQPDTSVVKDSQV 668
DQ 541 RPYKEMLLCIXVHDAGIVHSDLKPANFLVKGILKIIDFGIANAYPEHTVNIYRETQI 600
QY 669 GTVNYMPPEAIKDM-----SSRENGKSKSISPKSDVMSLGCILYIYMYTKTPFQQIINQ 724
DQ 601 GTPNYWAPALVAMNYTQNSENQHGNKWKVGRPSDMWSCGCIYQMIYKPPYGSFGQ 660
QY 725 ISKLHAIIDPNHETFFPD-----IPEKDIQDVLKCKLDRPKQISIPELLAHPYVQ 776
DQ 661 -NRLLAIMNPDKVIFPPEHTSNNKIP-KSAIELMKACLYRNPDRRTVDKVLSTFIQ 717

RESULT 5
TLK2 HUMAN
ID TLK2 HUMAN STANDARD; PRT; 772 AA.
AC Q86UE8; Q9UKI7; Q9Y4F7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase tousled-like 2 (EC 2.7.1.37) (Tousled-
DE like kinase 2) (PKU-alpha).
GN TLK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

```

[1] SEQUENCE FROM N.A. (ISOFORM 3), FUNCTION, SUBCELLULAR LOCATION, AND
RP TISSUE SPECIFICITY.
RX MEDLINE=98087437; PubMed=9427565;
RA Yamakawa A., Kameoka Y., Hashimoto K., Yoshitake Y., Nishikawa K.,
Tanihara K., Date T.;
RT "cDNA cloning and chromosomal mapping of genes encoding novel protein
kinases termed PKU-alpha and PKU-beta, which have nuclear localization
signal.";
RL Gene 202:193-201(1997).
[2]
RP SEQUENCE FROM N.A. (ISOFORM 2), FUNCTION, MUTAGENESIS OF ASP-592,
RP SUBCELLULAR LOCATION, INTERACTION WITH TLK1, AND ENZYME REGULATION.
RC TISSUE=Placenta;
RX MEDLINE=99452775; PubMed=10523312;
RA Sillje H.H.W., Takahashi K., Tanaka K., Van Houwe G., Nigg E.A.;
RT "Mammalian homologues of the plant tousled gene code for cell-cycle-
regulated kinases with maximal activities linked to ongoing DNA
replication.";
RL EMBO J. 18:5691-5702(1999).
[3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Strausberg R.D., Collins F.S., Wagner J., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Frange C.,
Raha S.S., Loquellanc N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettelman M., Maman A., Rodriguez S., Sanchez A.,
Whiting M., Maman A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
RP FUNCTION, AND ENZYME REGULATION.
RX MEDLINE=22546651; PubMed=12660173;
RA Groth A., Lukas J., Nigg E.A., Sillje H.H.W., Wernstedt C., Bartek J.,
Hansen K.;
RT "Human tousled like kinases are targeted by an ATM- and Chk1-dependent
DNA damage checkpoint.";
RL EMBO J. 22:1676-1687(2003).
CC -!- FUNCTION: Rapidly and transiently inhibited by phosphorylation
following the generation of DNA double-stranded breaks during S-
phase. This is cell cycle checkpoint and ATM-pathway dependent and
appears to regulate processes involved in chromatin
assembly.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- COFACTOR: Magnesium.
CC -!- ENZYME REGULATION: Cell-cycle regulated, maximal activity in S-
phase. Inactivated by phosphorylation at Ser-749, potentially by
CHK1.
CC -!- SUBUNIT: Heterodimerizes with TLK1.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q86UE8-1; Sequence=Displayed;
CC Note=No experimental confirmation available;
CC Name=2;
CC IsoId=Q86UE8-2; Sequence=VSP_050573;
CC Name=3;
```


STRAIN=C57BL/6J; TISSUE=Testis;
 MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Giasi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawaasa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lerhard B., Lyons P.A.,
 RA MacLott D.R., Maltakis L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Nomata K., Okido T., Pavan W.J., Partea G., Pesole G.,
 RA Perovskiy N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sancelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki K., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.,
 RA "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573 (2002).
 RP [3]
 RN TISSUE SPECIFICITY.
 RX MEDLINE=99452775; PubMed=10523312;
 RA Sillje H.H.W., Takahashi K., Tanaka K., Van Houwe G., Nigg E.A.,
 RA "Mammalian homologues of the plant touselid gene code for cell-cycle-
 RT regulated kinases with maximal activities linked to ongoing DNA
 RT replication";
 RL EMBO J. 18:5691-5702(1999).
 CC -!- FUNCTION: Rapidly and transiently inhibited by phosphorylation
 CC following the generation of DNA double-stranded breaks during S-
 CC phase. This is cell cycle checkpoint and ATM-pathway dependent and
 CC appears to regulate processes involved in chromatin assembly (By
 CC similarity).
 CC -!- FUNCTION: Testis-specific isoforms may play a role in
 CC spermatogenesis.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- COFACTOR: Magnesium (By similarity).
 CC -!- ENZYME REGULATION: Cell-cycle regulated, maximal activity in S-
 CC phase. Inactivated by phosphorylation at Ser-749, potentially by
 CC CHK1 (By similarity).
 CC -!- SUBUNIT: Heterodimerizes with TLK1 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Comment=Additional isoforms seem to exist;
 CC Name=1;
 CC IsoId=O55047-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=O55047-2; Sequence=VSP 050575, VSP 050576;
 CC Note=No experimental confirmation available;
 CC Name=3;
 CC IsoId=O55047-3; Sequence=VSP 050574;
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed in all tissues
 CC examined, with high levels in heart and testis, in particular the
 CC pachyene spermatocytes and in round spermatids. Some evidence for
 CC the existence of a testis-specific isoform suggesting a role in
 CC spermatogenesis.
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).

EMBL; AF045252; RAC02225.1; AUT_INIT.
 EMBL; AF045254; RAC02227.1; -
 EMBL; AF045253; RAC02226.1; -
 EMBL; AK014829; BAB29570.2; -
 MGD; MGI:1346023; TLK2.
 GO; GO:0005634; C:nucleus; ISS.
 GO; GO:0005524; F:ATP binding; ISS.
 GO; GO:0004674; F:protein serine/threonine kinase activity; ISS.
 GO; GO:0007242; P:intracellular signaling cascade; ISS.
 GO; GO:0006468; P:protein amino acid phosphorylation; ISS.
 GO; GO:0001672; P:regulation of chromatin assembly/diseassembly; ISS.
 InterPro; IPR000719; Prot kinase.
 InterPro; IPR008271; Ser_Thr_kinase.
 InterPro; IPR001290; Ser_thr_kinase.
 Pfam; PF00569; pkinase; 1.
 ProDom; PD000001; Prot kinase; 1.
 SMART; SM00220; S_TKc; 1.
 SMART; SM00219; TyKc; 1.
 PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KX Spermatogenesis; Chromatin regulator; ATP-binding;
 KW Serine/threonine-protein kinase; Transferase; Cell cycle; DNA damage;
 KW Nuclear protein; Coiled coil; Phosphorylation; Alternative splicing.
 FT DOMAIN 193 244
 FT DOMAIN 285 345
 FT DOMAIN 349 397
 FT DOMAIN 408 687
 FT NP_BIND 414 422
 FT BINDING 437 437
 FT ACT_SITE 538 538
 FT MOD_RES 696 696
 FT VARSPPLIC 2 117
 FT VARSPPLIC 89 89
 FT VARSPPLIC 626 718
 FT VARSPPLIC 626 718
 FT CONFLICT 520 520
 FT CONFLICT 525 525
 FT CONFLICT 535 535
 FT CONFLICT 540 540
 FT CONFLICT 543 543
 FT CONFLICT 548 548
 FT CONFLICT 555 555
 SQ SEQUENCE 718 AA; 82260 MW; 60009617C162BB05 CRC64;

Query Match 7.7%; Score 335; DB 1; Length 718;
 Best Local Similarity 22.1%; Pred. No. 5.8e-10;
 Matches 161; Conservative 128; Mismatches 268; Indels 172; Gaps 32;
 QY 155 ERGAVPLEMLEIALRNINLQKQLSEEEKKLSASTVLTAQSPFSGSLGHQLQNNNSCD 214
 DB 110 EQSALP-TLMSVNLAKPRLDTQLAPRGAG--LCFTFVSAQQNSPSTGS-GNTEHSCS 164
 QY 215 SRGOTTKARFLGCVNPPDADIGYNSLRQTKNQSCPCFGRVFNILNSPDCDVKTD- 273
 DB 165 SQQGIS-----IQHROTQSDLTIEKISALENSDLEKKEGRID-DLURA-NCDLRRQI 217
 QY 274 DSVVPCFMKQTSRSEC-----RDLVVPGSK-----PSGNDSCELRLKSVQN-SHPK 320
 DB 218 DEQCKMLEKVKELNCTVMSKKLLLEKSKQEKACRDKSMQDLRLGLHFTTVRHGASFT 277

QY 321 EPLVSDKSSLLITDSTLTKNKTSSLLAKLETKEYQDEVPESNQKQAKRSECI 380
 Db 278 EQW-----TDGAFON-----LIQOERINSOREEI--ERQKVLAKRKPAM 318
 QY 381 NQNPASSNHQIPELARKVTEQKHTTFEQPVSVSKQSPISSTKWFDPKICKTPSS 440
 Db 319 QOAPPATNEQK-----RKSTN-----GAENETLTAETHOEELFKLRLG 360
 QY 441 NTLDDYMSCFRTPVVKNDPPACQLSTPYGQPCFOQOQHQLATPLQNTQVLASSANE 500
 Db 361 H-----LKEEAEIQA-ELERLERNLHIRE 386
 QY 501 CISVGR-----IYSILKQIGSGGSKVQVLEKQIY-AIKYNLEEA-D 545
 Db 387 LKRIHEDNSQFKDHPITLNDRYLLHLGRGFSVYKAFDLTEQRYVAVKIHLQNKWR 446
 QY 546 NQTLDSYNEIAYLNKLQOHSN--KIIRLYDYEI--TDQIYVMWECGNIDLSNWLKKKK 601
 Db 447 DEKENYHKAACREYRIHKELDHPRIVKLYDFSLDTSFCTVLEYCEGNDLDFYLKQHK 506
 QY 602 SIDWERKSYWNMLEAVHTIQ--HGIVHSDLKPNFLIVDM-----LKLIDFGIANQM 655
 Db 507 LMSEKARSITMQIVNALKYLNKPIIHYDLKPGNILLWNGTACGEIKITDFGLSKIM 566
 QY 656 QPDTTSVVKDSQV-----GTVMPPPEAIKDMSSSRENGSKSKISPKSDWSLGCILYY 710
 Db 567 DDDSYNSVDGMLTSQAGATWYLPPECF-----VVGEPKLSNKVDVWSGVIFYQ 619
 QY 711 MYQKTFP-----QIINQISKULHAIIDNHEIEFPDIP--EKDLOVLKCLKRPDPK 762
 Db 620 CLYGRKPGHNSQOQILQENTILKAT-----EVQFPKPVVPEAKAFIRRCCLAYRKED 674
 QY 763 RISIPELLAHYVVOIQTHPVNQMAKGTTEEMKYVLGVLGINSNSLTKAAKTYLHYSG 822
 Db 675 RIDVQQLACDYP--LLPH-----IRKSVTSPAGAAIAS-----TS 709
 QY 823 GESHSNSSS 831
 Db 710 GASNSSSN 718

RESULT 7

TLK1_HUMAN
 ID TLK1_HUMAN STANDARD; PRT; 766 AA.
 AC Q9UKI8; Q14150; Q8N591; Q9NYH2; Q9Y4F6;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Serine/threonine-protein kinase tousled-like 1 (EC 2.7.1.37) (Tousled-like kinase 1) (PKU-beta).
 GN TLK1 OR KIAA0137.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2), FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
 RC TISSUE=Placenta, and Testis;
 RX MEDLINE=98087437; PubMed=9427565;
 RA Yanakawa A., Kameoka Y., Hashimoto K., Yoshitake Y., Nishikawa K., Tanihara K., Date T.;
 RT "cDNA cloning and chromosomal mapping of genes encoding novel protein kinases termed PKU-alpha and PKU-beta, which have nuclear localization signal.";
 RT signal.";
 RL Gene 202:193-201(1997).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, MUTAGENESIS OF ASP-586. SUBCELLULAR LOCATION, INTERACTION WITH TLK2, AND ENZYME REGULATION.
 RC TISSUE=Placenta;
 RX MEDLINE=99452775; PubMed=10523312;
 RA Sillje H.H.W., Takahashi K., Tanaka K., Van Houwe G., Nigg E.A.;
 RT "Mammalian homologues of the plant tousled gene code for cell-cycle-

RT regulated kinases with maximal activities linked to ongoing DNA replication.";
 RL EMBO J. 18:5691-5702(1999).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 3), FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
 RC TISSUE=Placenta;
 RX MEDLINE=20056088; PubMed=10588641;
 RA Cabaniols J.-P., Ravichandran V., Roche P.A.;
 RT "Phosphorylation of SNAP-23 by the novel kinase SNAK regulates t-SNARE complex assembly.";
 RT Mol. Biol. Cell 10:4033-4041(1999).
 RL [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Bone marrow;
 RX MEDLINE=96127530; PubMed=8590280;
 RA Nagase T., Seki N., Tanaka A., Iehikawa K.-I., Nomura N.;
 RT "Prediction of the coding sequences of unidentified human genes. IV. The coding sequences of 40 new genes (K1AA0121-K1AA0160) deduced by analysis of cDNA clones from human cell line KG-1.";
 RL DNA Res. 2:1167-1174(1995).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=uterus;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F., Diatchenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshitaki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [6]
 RP FUNCTION, MUTAGENESIS OF SER-743, PHOSPHORYLATION OF SER-743, AND ENZYME REGULATION.
 RX MEDLINE=22546651; PubMed=12660173;
 RA Groth A., Lukas J., Nigg E.A., Sillje H.H.W., Wernstedt C., Bartek J., Hansen K.;
 RT "Human Tousled like kinases are targeted by an ATM- and Chk1-dependent DNA damage checkpoint.";
 RL EMBO J. 22:1676-1687(2003).
 CC -!- FUNCTION: Rapidly and transiently inhibited by phosphorylation following the generation of DNA double-stranded breaks during S-phase. This is cell cycle checkpoint and ATM-pathway dependent and appears to regulate processes involved in chromatin assembly. Isoform 3 phosphorylates and enhances the stability of the t-SNARE SNAP23, augmenting its assembly with syntaxin.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- COFACTOR: Magnesium.
 CC -!- ENZYME REGULATION: Cell-cycle regulated, maximal activity in S-phase. Inactivated by phosphorylation at Ser-743, potentially by Chk1.
 CC -!- SUBUNIT: Heterodimerizes with TLK2.
 CC -!- SUBCELLULAR LOCATION: Nuclear. Isoform 3 appears to be cytoplasmic.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=Q9UKI8-1; Sequence=Displayed;
 CC Name=2;

CC isoid=Q9UK18-2; Sequence=VSP_050571;
 CC Name=3; Synonym=SNAK;
 CC isoid=Q9UK18-3; Sequence=VSP_050570;
 CC -!- TISSUE SPECIFICITY: Widely expressed. Present in fetal placenta,
 CC liver, kidney and pancreas but not heart or skeletal muscle. Also
 CC found in adult cell lines. Isoform 3 is ubiquitously expressed in
 CC all tissues examined.
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB004885; BAA20562.1; -;
 CC EMBL; AF162666; AAF03094.1; ALT_INIT.
 CC EMBL; AF246219; AAF71263.1; -;
 CC EMBL; D50927; BAA09486.2; ALT_INIT.
 CC EMBL; BC032657; AAH32657.1; -;
 CC Genew; HGNC:11841; TLK1
 CC GO; GO:0005634; C:nucleus; IE:protein serine/threonine kinase activity; IDA.
 CC GO; GO:0005524; F:ATP binding; IDA.
 CC GO; GO:0004674; P:intracellular protein transport; IDA.
 CC GO; GO:0006886; P:intracellular signaling cascade; IDA.
 CC GO; GO:0007242; P:intracellular amino acid phosphorylation; IDA.
 CC GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
 CC GO; GO:0001672; P:regulation of chromatin assembly/disassembly; IDA.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR008271; Ser_thr_pkin_AS.
 CC InterPro; IPR002290; Ser_thr_pkinase.
 CC InterPro; IPR001245; Tyr_pkinase.
 CC Pfam; PF00069; pkinase; 1.
 CC ProDom; PD000001; Prot_kinase; 1.
 CC SMART; SM00220; S_TKc; 1.
 CC SMART; SM00219; TyrKc; 1.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 CC Chromatin regulator; ATP-binding; Serine/threonine-protein kinase;
 KW Transferrase; Cell cycle; DNA damage; Nuclear protein; Coiled coil;
 KW Phosphorylation; Alternative splicing
 FT DOMAIN 230 281
 FT DOMAIN 397 445
 FT DOMAIN 456 734
 FT NP_BIND 462 470
 FT BINDING 485 485
 FT ACT_SITE 586 586
 FT MOD_RES 743 743
 FT VARSPPLIC 1 217
 FT VARSPPLIC 136 136
 FT MUTAGEN 586 586
 FT MUTAGEN 743 743
 FT MUTAGEN 743 743
 FT MUTAGEN 743 743
 FT MUTAGEN 743 743
 FT CONFLICT 88 88
 FT CONFLICT 102 102
 FT CONFLICT 230 230
 FT CONFLICT 261 261
 FT CONFLICT 416 416
 FT CONFLICT 439 439
 FT CONFLICT 471 471
 FT CONFLICT 477 477
 FT CONFLICT 525 525
 FT CONFLICT 525 525
 FT CONFLICT 665 665

2).
 /FTid=VSP_050571.
 D->A: LOSS OF KINASE ACTIVITY.
 S->A: LOSS OF KINASE INHIBITION IN
 RESPONSE TO DNA DAMAGE.
 S->E: LOSS OF KINASE INHIBITION IN
 RESPONSE TO DNA DAMAGE.
 S->D: LOSS OF KINASE INHIBITION IN
 RESPONSE TO DNA DAMAGE.
 S->T (IN REF. 2).
 G->E (IN REF. 1).
 Q->L (IN REF. 1).
 E->D (IN REF. 1).
 N->G (IN REF. 1).
 E->H (IN REF. 1).
 Y->D (IN REF. 1).
 Y->S (IN REF. 1).
 D->V (IN REF. 1).
 F->Y (IN REF. 1).

FT CONFLICT 730 730 N -> C (IN REF. 1).
 SQ SEQUENCE 766 AA; 86699 MW; A65AE0A2A7C7FF2F CRC64;
 Query Match 7.6%; Score 332; DB 1; Length 766;
 Best Local Similarity 21.3%; Pred. No. 8.9e-10;
 Matches 171; Conservative 124; Mismatches 268; Indels 238; Gaps 34;
 QY 132 SFAQFELS--QCNVKKSKQLQKAVRGAVPLMELEIALRNINLQKOLLKOLLSBEKKNLISA 189
 DB 15 SWSQLSTSTPSSAAARSLLNHTPPSGR--PREGAMDELHSLDPRQELLEARFTGVASG 73
 QY 190 STVLT-----AOESFGSLGHLQNRNN-----SCDSRGQTTKARFLYG 227
 DB 74 STGTSGSCSVGAKASTNNESHSP--GSLGSLSDKSESTPEKKQSESGRRKAE--N 129
 QY 228 ENMPDQDAEIGVYRSLRQNTKTKQSCFGRVFNLLNGPDCVKTDDSVVPCEMKRQTSR 287
 DB 130 QNESSQKSGIGRGH-----KISD-----YFYQ--- 153
 QY 288 SECDLWVEGSKPSGNDSCELRN-----LKSVQNSHF-----KEP 322
 DB 154 -----GGNGSPVRGIPPAIRSPONSHSHSTPSSVRPNPSPPTALAFGHP 200
 QY 323 LYSDEK--SSELIITDSITLKNKTESLAKLEETK-----EYQEPVP----- 364
 DB 201 IVQPKLSPKIITDITMLK-----LALESNKIQLEKKEGRIIDLLRANCDLRQI 253
 QY 365 ESNQK--QWQAKRKSECN-----LALESNKIQLEKKEGRIIDLLRANCDLRQI 388
 DB 254 DQCKLEKLYKELNKCISMSKLLIEKSTQBLEKSSREKSMQDLRLGLHFTTVRHGASPT 313
 QY 389 NHW-----QIPELARKVN--TEQHTTFEQDFVSQSPISSTKSWFDPKSIKCTSSNTL 443
 DB 314 EQMTDGFAPQNLVKKQEWNVNQOREDIERQKLAKRKPTANNS-----QAPSTNSE 365
 QY 444 DDVM-----SCFRTPVKNPDPACQLSTPYGQACFO-----QQQHILATPLQNLQ 491
 DB 366 PKRKNKAVNGAENDPFFVRPNLPOLLTLAEYHEQBEIFKLRLGHLKKEAEIQAELERL 425
 QY 492 VIASSANECISVK-----GRIYILKQIGSGSKSVQVQLNEKKQIY--ATK 537
 DB 426 RVNHLHIRELKRINNEEDNSQFKDHTNLERYLLHLGREGFSEYKAFDLYEQRYAAVK 485
 QY 538 VYNLEEA--NQDTLDSYNEIAYNLKQCHSD--KIIRLYDYEI--TDQVIYVMECGNID 592
 DB 486 IHLNKSWEDEKENHAKREYRHKELDHPRIKLYDYFSLDITDFTCTVLECEGND 545
 QY 593 LNSWLKSKSIDPWEKSKYKMLKLEAVHTIQ--HGIVHSDLKPANFLIVDM---LKL 646
 DB 546 LDFYLLQKHLKSEKEARSIVMQLVNALRYLNEIKPPIIHYDLKPGNILLVDGTACGEIKI 605
 QY 647 IDFGIANQMPDTTSV-----VKDSQVGTVMYVPEPAIKDMSSRENGSKSKISPKSDYV 702
 DB 606 TDFGLSKIMDDDSYGVDMGLTSSQAGTYWYLPEPCF-----VVGKEPPKISNKVDVW 658
 QY 703 SLGCILYVMTYGTPTF-----QQINQISKLHAIIDPNHEIEFPDIP--EKDLQDLVK 754
 DB 659 SVGVIFQCLYGRKPGHNSQOQDILQENTLKAT-----EVQFPVKPVVSSEAKAFIR 713
 QY 755 CLKRDPKQRISEIPILLAHPIV 775
 DB 714 CLAYRDEDFVHQLANDPYL 734
 RESULT 8
 TKL1_MOUSE
 ID TKL1_MOUSE STANDARD; PRT; 766 AA.
 AC Q8C0V0;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Serine/threonine-protein kinase tousled-like 1 (EC 2.7.1.37) (Tousled-
 DE like kinase 1).


```

QY 293 LVVFGSKSGNDSCELRNLKSVQNHFKPEPLVSDKESGSELI 334
Db 258 KLEKYEKRLN-KCISMSKLLIBKSTQEKLSRSKQMDRLRLGHFTTVRHGASFTEQW 316
QY 335 TDSITLKNKTESLILAKLEETKEYQEEPEVPESNOKQWAKRKSCINONPAASNNHWQIP 394
Db 317 TDGFAFN-----LVKQEWNVQREDI--ERQKLLGKKEKPTANNSSQAPATN----- 363
QY 395 ELARKVTEQKHTTFBQPVFSVSKSPPISTSKWFDPKSICKTPSSNLTDDYMSCFPTPV 454
Db 364 -----SEAKQR-----KTKAVNGAEN-----DPF 382
QY 455 VKNDFPPACQLSTPYGOPACFQ-----QOOHQILATPLQNLQVLASSANECSIVK-- 505
Db 383 VRPNLPQLLTIAEYHESQEEIFKLLGHLKKEEARIQELERLERNLHRELKRNED 442
QY 506 -----CRIVSILKQIGSGSSKVQVLNKKQIY-AIKYVNLEEA-DNQTLDSYR 553
Db 443 NSQFKDPTLNERNYLLHLLGRGGFSEVYKAFDLYEQRYAAVKIHLQNLKSWRDEKENYH 502
QY 554 NEIAYLNKLQOHS-D-KIIRLYDYEI--TDQYIYMWECGNIDLSNLWKKKSIDPWPERK 609
Db 503 KHACREYRIKELDHPRIVKLYDFSLDTDTCTVLSYCEGNDLDFYLRQHLKMSKEAR 562
QY 610 SYWQMLAEVHTIHQ--HGIYHSLDKPANELIVDGM-----LKLIDFGIANQMOPDITSV- 662
Db 563 SINWQIVNALRYNEIKPPIIHYDLKFGNILLVDGTACGEIKITDFGLSKIMDDDSYGVND 622
QY 663 ---VKDSQGVYVNWMPKAKDMSSRENGSKSKISPKSDVWSLGSILYIMYTKPTP- 718
Db 623 GMDLTSQAGFYWLPPCECF-----VWGKEPKKSNKVDVMSVGVIFQCILYGRKPGF 675
QY 719 -----QOILINGISKLHAIIDPNHEIFPDIP--EKDLQDVLKCLKRDPKORISIPELLA 771
Db 676 HNSQCDILQENTILKAT-----EVQPFVFPVSSEAKAFIRCLAYRKDRDFVHQLAN 730
QY 772 HPYV 775
Db 731 DPYL 734

RESULT 9
CDS1_SCHPO STANDARD; PRT; 460 AA.
AC Q93170; O42642;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein kinase cds1 (EC 2.7.1.37) (Checkpoint kinase cds1).
GN CDS1 OR SPCC18B5.11C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=95240713; PubMed=7723827;
RA Murakami H.; Okayama H.;
RT "A kinase from fission yeast responsible for blocking mitosis in S
phase.";
RL Nature 374:817-819 (1995).
RN [2]
RP SEQUENCE FROM N.A., FUNCTION, SUBUNIT, AND PHOSPHORYLATION.
RX STRAIN=972;
RX MEDLINE=98119835; PubMed=9450932;
RA Lindsay H.D., Griffiths D.J.F., Edwards R.J., Christensen P.U.,
RA Murray J.M., Osman F., Walworth N., Carr A.M.;
RT "S-phase-specific activation of Cds1 kinase defines a subpathway of
the checkpoint response in Schizosaccharomyces pombe.";
RL Genes Dev. 12:382-395 (1998).
RN [3]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Guillermo R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feitwell T., Fraser A.,
RA Gentles S., Goble N., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., Mclean J.,
RA Mooney P., Moutle S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellon J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voicakert G., Aert R., Robben J., Grymonprez B.,
RA Weijtens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880 (2002).
CC -!- FUNCTION: Has a role in the DNA replication-monitoring S/G2
CC checkpoint system. It is responsible for blocking mitosis in the
CC S phase. It monitors DNA synthesis by interacting with DNA
CC polymerase alpha and sends a signal to block the onset of mitosis
CC while DNA synthesis is in progress.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBUNIT: Interacts with rad26.
CC -!- PTM: Autophosphorylated.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. CDS1
CC subfamily.
CC -!- SIMILARITY: Contains 1 FHA domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X85040; CAA59410.1; -.
CC EMBL; AJ222869; CAA11019.1; -.
CC EMBL; AL109736; CAB52158.1; -.
CC FIR; S58882; S58882.
CC HSSP; Q00534; 1B17.
CC GenesDB SPombe; SPCC18B5.11c; -.
CC InterPro; IPR000253; FHA.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_thr_kin_AS.
CC InterPro; IPR002290; Ser_thr_kinase.
CC InterPro; IPR008984; SMAD FHA.
CC InterPro; IPR001245; Tyr_kinase.
CC Pfam; PF00498; FHA; 1.
CC Pfam; PF00669; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00240; FHA; 1.
CC SMART; SM00220; S_TK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00006; FHA_DOMAIN; 1.
CC TRANSFERase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Phosphorylation.
FT DOMAIN 60 116

```

James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B., Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Leinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B., Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Sipakovski G.V., Ussey D., Barrell B.G., Nurse P., "The genome sequence of *Schizosaccharomyces pombe*." Nature 415:871-880 (2002).

-!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNF1 subfamily.

-!- SIMILARITY: Contains 1 UBA domain.

--- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

--- EMBL: AL031543; CAA20833.1; --
 PIR: T41587; T41587.
 HSP: Q63450; IA06.
 GeneDB Spombe; SPCC74.03c; --
 InterPro: IPR000719; Prot_kinase.
 InterPro: IPR008271; Ser_thr_pkin_AS.
 InterPro: IPR002290; Ser_thr_pkinase.
 InterPro: IPR000449; UBA_domain.
 Pfam: PF00069; pkinase_1.
 ProDom: PD000001; Prot_kinase; 1.
 SMART: SM00220; S_TKC; 1.
 PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 PROSITE: PS50030; UBA; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 34 285 PROTEIN KINASE.
 FT DOMAIN 304 345 UBA.
 FT NP_BIND 40 48 ATP (BY SIMILARITY).
 FT BINDING 63 63 ATP (BY SIMILARITY).
 FT ACT_SITE 156 156 BY SIMILARITY.
 SQ SEQUENCE 576 AA; 65996 MW; E5857E8F171E7B50 CRC64;

Query Match 7.4%; Score 324.5; DB 1; Length 576;
 Best Local Similarity 29.3%; Pred. No. 1.5e-09;
 Matches 102; Conservative 67; Mismatches 132; Indels 47; Gaps 15;

QY 509 YSILKQIGSGSGSKVQVLNEK-KQIYAIKYVN---LEADNQTLDSYRNEYATLKLQ 564
 DB 34 YIIEITLGEFGFKVLATHYTKQVKALFKFISRLQLKSDMHM-RVEREISYL-KLLR 90
 QY 565 HSDKIIRLYDIYEITDQYIYVWMECGNIDLSNWLKKKSIDPWERKSYKWMLEAVHTIQ 624
 DB 91 H-PHIILKLYDVIITPTDVIWVIEYAGGELFDYIEKKRMTDEGRFFQIICALEYCHR 149
 QY 625 HGIVHSDLPKPNFLIVDGM-LKILDFGIANQOPDTTSVVXKDSQGVTVNMPPEAKDMS 683
 DB 150 HKIVHRDLKENLLDDNLNVLKIDAGLSNIM---TDGNFLTSCGSPNYAAPEVI--- 202
 QY 684 SSRENGSKSKISPKSDVMSLGLCIIYMYTGYKTPFQQ---IINQISKLHAIDPNHIEFP 741

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B., Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Leinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B., Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Sipakovski G.V., Ussey D., Barrell B.G., Nurse P., "The genome sequence of *Schizosaccharomyces pombe*." Nature 415:871-880 (2002).

-!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNF1 subfamily.

-!- SIMILARITY: Contains 1 UBA domain.

--- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

--- EMBL: AL031543; CAA20833.1; --
 PIR: T41587; T41587.
 HSP: Q63450; IA06.
 GeneDB Spombe; SPCC74.03c; --
 InterPro: IPR000719; Prot_kinase.
 InterPro: IPR008271; Ser_thr_pkin_AS.
 InterPro: IPR002290; Ser_thr_pkinase.
 InterPro: IPR000449; UBA_domain.
 Pfam: PF00069; pkinase_1.
 ProDom: PD000001; Prot_kinase; 1.
 SMART: SM00220; S_TKC; 1.
 PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 PROSITE: PS50030; UBA; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 34 285 PROTEIN KINASE.
 FT DOMAIN 304 345 UBA.
 FT NP_BIND 40 48 ATP (BY SIMILARITY).
 FT BINDING 63 63 ATP (BY SIMILARITY).
 FT ACT_SITE 156 156 BY SIMILARITY.
 SQ SEQUENCE 576 AA; 65996 MW; E5857E8F171E7B50 CRC64;

Query Match 7.4%; Score 324.5; DB 1; Length 576;
 Best Local Similarity 29.3%; Pred. No. 1.5e-09;
 Matches 102; Conservative 67; Mismatches 132; Indels 47; Gaps 15;

QY 509 YSILKQIGSGSGSKVQVLNEK-KQIYAIKYVN---LEADNQTLDSYRNEYATLKLQ 564
 DB 34 YIIEITLGEFGFKVLATHYTKQVKALFKFISRLQLKSDMHM-RVEREISYL-KLLR 90
 QY 565 HSDKIIRLYDIYEITDQYIYVWMECGNIDLSNWLKKKSIDPWERKSYKWMLEAVHTIQ 624
 DB 91 H-PHIILKLYDVIITPTDVIWVIEYAGGELFDYIEKKRMTDEGRFFQIICALEYCHR 149
 QY 625 HGIVHSDLPKPNFLIVDGM-LKILDFGIANQOPDTTSVVXKDSQGVTVNMPPEAKDMS 683
 DB 150 HKIVHRDLKENLLDDNLNVLKIDAGLSNIM---TDGNFLTSCGSPNYAAPEVI--- 202
 QY 684 SSRENGSKSKISPKSDVMSLGLCIIYMYTGYKTPFQQ---IINQISKLHAIDPNHIEFP 741

Thu Sep 30 10:17:13 2004

```
Ds 1581 LESDPKRWNTASQLLDHSFVKVCT 1604

RESULT 12
YKT3 CAEEL STANDARD; PRT; 1231 AA.
AC P34314; P34323;
DT 01-FEB-1994 (Rel. 28, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative serine/threonine-protein kinase C07A9.3 in chromosome III
DE (EC 2.7.1.-).
DE C07A9.3.
GN Caenorhabditis elegans.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peridodinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Bristol N2; PubMed=7906398;
RX MEDLINE=94150718;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL elegans.";
RL Nature 368:32-38 (1994).
RN [2]
RP REVISIONS.
RA Durbin R.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z29094; CAA82347.2; -
DR PIR; S40704; S40704.
DR PIR; S40712; S40712.
DR WormPep; C07A9.3; CE20487.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR InterPro; IPR000290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase_1.
DR ProDom; PDC00001; Prot_kinase; 1.
DR SMART; SMO0220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW ATP-binding.
FT DOMAIN 423 496 GLN-RICH.
FT DOMAIN 922 1199 PROTEIN KINASE.
FT NP_BIND 928 936 ATP (BY SIMILARITY).
FT BINDING 951 951 ATP (BY SIMILARITY).
FT ACT_SITE 1052 1052 BY SIMILARITY.
SQ SEQUENCE 1231 AA; 139865 MW; 97C3204FAD90D95D CRC64;

Query Match 7.1%; Score 309; DB 1; Length 1231;
Best Local Similarity 20.1%; Pred. No. 2.3e-08;
```

OS Schizosaccharomyces pombe (Fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 CC NCBI TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90226892; PubMed=9321395;
 RA Samejima I., Mackie S., Fantes P.A.;
 RT "Multiple modes of activation of the stress-responsive MAP kinase
 RT pathway in fission yeast."
 RL EMBO J. 16:6162-6170(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., William R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgroves J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Gobie A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Moorey P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Symmonds R., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Weijens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe."
 RL Nature 415:871-880(2002).
 RN [3]
 RP SEQUENCE OF 96-1401 FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=97282620; PubMed=9136929;
 RA Shieh J.C., Wilkinson M., Buck V., Morgan B., Makino K.,
 RA Millar J.B.A.;
 RT "The Mcs4 response regulator coordinately controls the stress-
 RT activated Waki1-Wis1-Sty1 MAP kinase pathway and fission yeast cell
 RT cycle."
 RL Genes Dev. 11:1008-1022(1997).
 RN [4]
 RP SEQUENCE OF 457-543 FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=97331762; PubMed=9188094;
 RA Shiozaki K., Shiozaki M., Russell P.;
 RT "Mcs4 mitotic catastrophe suppressor regulates the fission yeast cell
 RT cycle through the Wki1-Wis1-Spc1 kinase cascade."
 RL Mol. Biol. Cell 8:409-419(1997).
 CC -1- FUNCTION: Involved in a signal transduction pathway that is
 CC activated in under conditions of heat shock, oxidative stress or
 CC limited nutrition. Unlike wki1, it is not activated by changes in
 CC the osmolality of the extracellular environment. Activates the
 CC wsi1 MAP kinase kinase by phosphorylation.
 CC -1- SIMILARITY: Belongs to the Ser/thr family of protein kinases. MAP
 CC kinase kinase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities is prohibited.

CC entities requires a license agreement (see <http://www.iab-sib.ch/announcement>)
 CC or send an email to license@sb-sib.ch.
 CC -----
 DR EMBL; Z98763; CAB11500.1; -;
 DR EMBL; Y07750; CAA69030.1; ALT_INIT.
 DR EMBL; Y11989; CAA72718.1; -;
 DR EMBL; U81521; AAB39762.1; -;
 DR FIR; T39225; T39225.
 DR HSSP; P24941; 1CKP.
 DR GeneDB SPombe; SPAC9G1.02; -;
 DR GO; GO:0007049; P:cell cycle; ISS.
 DR GO; GO:000165; P:MAPKK cascade; ISS.
 DR GO; GO:0007165; P:signal transduction; ISS.
 DR InterPro; IPRO00719; Prot_kinase.
 DR InterPro; IPRO08271; Ser_Chk_kin_AS.
 DR InterPro; IPRO02290; Ser_Chk_pkinase.
 DR InterPro; IPRO01245; Tyr_kinase.
 DR Pfam; PF00669; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC_1
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Transferrase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 1037 1306 PROTEIN_KINASE.
 FT NP_BIND 1043 1051 ATP (BY SIMILARITY).
 FT BINDING 1066 1066 ATP (BY SIMILARITY).
 FT ACT_SITE 1161 1161 BY SIMILARITY.
 FT CONFLICT 484 485 RL -> SP (IN REF. 3).
 FT CONFLICT 1346 1346 D -> V (IN REF. 1).
 SQ SEQUENCE 1401 AA; 160537 MW; 529A989DB627F3D CRC64;

 Query Match 7.0%; Score 304.5; DB 1; Length 1401;
 Best Local Similarity 22.1%; Pred. No. 4.5e-08;
 Matches 129; Conservative 86; Mismatches 199; Indels 171; Gaps 22;

 Qy 304 DSEELRLNKS-----VQNSHKEPLVSDKSESLIITDSITLKVK----- 343
 Db 887 DECDLMDRKTFRGVGALEFLMLMI--RGNNLLIIDDAMFLKIREKVGKSMAFLLTHFDV 944
 Qy 344 --TESSLAKLEETKEYQEEVPESNQKWQAKRKSECINQNPAAASNHWQIPELARKVN 401
 Db 945 LGAKSKVAKLQ-----RESTEVSSPLTS-FGDVEEALSQ 982
 Qy 402 TEQKHTTFEQPVSVSKQSPPISTKSWFDKPKCTKTPSSNTLDDYMSCTPTPVVKNDDPP 461
 Db 983 LLQKETMLRIDELEIERNNLTLE-----RLAIGHVLDD--SVFR---NRDF-- 1023
 Qy 462 ACOLSTPYGQPAQFOQHQHQLATPLQLQLVLASSANECISVKGRYISILKQIGSGSS 521
 Db 1024 -IKLASSFSNITRWQOCH-----FVRSQWFG 1049
 Qy 522 KVFQVLN-EKKQIYAIKYNLEAD--NQTLDSYRNEIAVLNKLOQHSDKIIRLYDYVEIT 578
 Db 1050 DVTGTVMNETGDLIAVKEIKLQDSRTFRSDVDQIHNETMVLRL--NHPNVVTYGVGEVH 1107
 Qy 579 DQVIYMYWE-CGNIDILNSWLKKKSDIPWERKSYWKNMLEAVHTIQHGIVHSDLRKANE 637
 Db 1108 REKVIYFMFCQGGSLADLLAHGRIEDENVLYKVYVQLLEGSLAYIHSQHLHREIKPANI 1167
 Qy 638 LIVD-GMLKUIDFGIANQMPDITTSVVKQSQV-----GTVNYMPEPAIKDMSSREN 688
 Db 1168 LLDRHGMIKYSDFGSALYVSFPDTPDEVRYEDIQPELQHLAGTPTWYMAPEIL----- 1219
 Qy 689 GKSKSIKPSKSDVWSLGCILYVMTYGTPTFQQLINOISKLHAIIDPNHETEFDPDIPEKD- 747
 Db 1220 GTKKGDFFG-AMDVLSLGCIVLEMMTGSTPSEMDNEWAINYHVA---AHTSISIQNEK 1274
 Qy 748 ----LDQVLKCCLRKDPKQRISSIPELLAHYV-----QIQTHPVNQAKGTT----- 790
 Db 1275 ISSLARDFIQCCFERDPEQRPRAVDLLTHPWIIDFRKKTIITWPPATITKTSLSHTITE 1334

QY 791 -----BEMKYVLGQLVGLNSPNSI 809
Db 1335 EKTAQLAGRHDDSKAETSLAASYKEESALPVASVGLRQPNEL 1379
RESULT 14
M3K4 MOUSE
ID M3K4 MOUSE STANDARD; PRT; 1597 AA.
AC O08648; O08649; O70124;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase 4 (EC 2.7.1.-)
DE (MAPK/ERK kinase kinase 4) (MEK kinase 4) (MEKK 4).
GN MAP3K4 OR MEKK4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A. (ISOFORMS A AND B).
RP TISSUE=Brain;
RX MEDLINE=97236778; PubMed=9079650;
RA Gewins P., Blank J.L., Johnson G.L.;
RT "Cloning of a novel mitogen-activated protein kinase kinase,
RT MEKK4, that selectively regulates the c-Jun amino terminal kinase
RT pathway.";
RL J. Biol. Chem. 272:8288-8295 (1997).
[2]
SEQUENCE OF 363-1049 FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Ectoplacental cone;
RX MEDLINE=97422608; PubMed=9268631;
RA Schweifer N., Valk P.J., Delwel R., Cox R., Francis F.,
RT Meier-Ewert S., Lehrach H., Barlow D.P.;
RT "Characterization of the C3 YAC contig from proximal mouse chromosome
RT 17 and analysis of allelic expression of genes flanking the imprinted
RT Igf2r gene.";
RL Genomics 43:285-297 (1997).
CC -!- FUNCTION: ACTIVATES THE CSBP2, P38 AND JNK MAPK PATHWAYS, BUT NOT
CC THE ERK PATHWAY. SPECIFICALLY PHOSPHORYLATES AND ACTIVATES MAP2K4
CC AND MAP2K6.
CC -!- SUBCELLULAR LOCATION: LOCALIZED IN PERINUCLEAR VESICULAR-LIKE
CC STRUCTURES, PROBABLY GOLGI-ASSOCIATED VESICLES.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=A;
CC IsoId=O08648-1; Sequence=Displayed;
CC Name=B;
CC IsoId=O08648-2; Sequence=VSP_004885;
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. HIGH EXPRESSION WAS FOUND IN
CC SKELETAL MUSCLE, KIDNEY, TESTIS FOLLOWED BY HEART BRAIN AND LUNG.
CC LOW EXPRESSION WAS FOUND IN SPLEEN.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. MAP
CC kinase kinase subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

DR EMBL; U85607; AAC53126.1; -;
DR EMBL; U85608; AAC53127.1; -;
DR EMBL; U66240; AAC08286.1; -;
DR HSP; P24941; LHCL;
DR MGD; MGI:1346875; Map3k4.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.

DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 1178 1182 POLY-ALA.
FT DOMAIN 1332 1390 PROTEIN KINASE.
FT NP_BIND 1338 1346 ATP (BY SIMILARITY).
FT BINDING 1361 1361 ATP (BY SIMILARITY).
FT ACT_SITE 1452 1452 BY SIMILARITY.
FT VARSPPLIC 1162 1213 Missing (in isoform B).
FT /FTid=VSP_004885.
FT CONFLICT 363 364 SL -> NS (IN REF. 2).
FT CONFLICT 473 473 T -> A (IN REF. 2).
SQ SEQUENCE 1597 AA; 179948 MW; E84AEAAE92D103A4 CRC64;
Query Match 6.9%; Score 303.5; DB 1; Length 1597;
Best Local Similarity 22.2%; Pred. No. 6e-08;
Matches 113; Conservative 99; Mismatches 165; Indels 131; Gaps 23;
QY 296 PCSKPSGNDSCSELNLSKVSQNSHFKPELVSDKSELIITDSTLKNKTSSLLAKLE-- 353
Db 1194 PG--PGGDSVPAPKPVNT-----APDTREGSVSPENDRL-----ASTAAELQFR 1234
QY 354 -----ETKEYQPEVPESN-----OKWOAKKSEKINQNPAASSNHWQIPELARKVN 401
Db 1235 SLRHSSTPERDEFPAYPRSDSSGSTRSEWELR---TLISQKDSASKQGPPIAIOKSVR 1291
QY 402 TEQKHTTPEQPVFVSQSPPISTSKMFPDKSIKCTPSSNTLDDYMSCTFTPVKNDPPP 461
Db 1292 -----LFEERRYREMRKNIIG-----QVCDTPKS--YDNVHVGLRKVT----- 1329
QY 462 ACQLSTPVGQACQOQOHOILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSS 521
Db 1330 -----FWQRG-----NKLGEGQYG 1344
QY 522 KVFQVLN-EKKQIYAIKYNLEADNQLDSYRNEIAYLNKLQSHSKIIRLDYDEITDQ 580
Db 1345 KVTCTISVDTGELMAMKEIRFPQNDHKTIKETADELKIFEGIKH--ENLVRYFGVELHRE 1402
QY 581 YIYVME-C--GNIDLSNLWKKKSIDPWERKSYKQWMLKLEAVHTIHOHGVHSDLKPAN- 636
Db 1403 EMYIFMEYCDTEGLLEEVSRGLQEHV-----IRLYTKQITVAINVLHGHVHRDIKGANI 1458
QY 637 FLIVDMLKLIIDFGIANOMQPDTTTSVVKD--SQVGTVNYMPPEAIKDMSSSRENGKSKK 694
Db 1459 FLTSSGLIKLGFCSVKLNKNAQTMPGEVNSTLGTAAAYMAPEVI-----TRAKGEGHGR 1513
QY 695 ISPKSDVNSLGLCIYYWYTKTPTFCQIINQISKLHAIDPNEHEFFDPIEK---DLQDV 751
Db 1514 ---AADIWSLGCWVEMVTGKRPWHEYEHNFQIMYK-VGMGHK---PPIPERLSPEGKAF 1566
QY 752 LKCCLRDPKORISIPPELLAHPPVQIOT 779
Db 1567 LSHCLSDPKIRWTASQJLLDAFVKVCT 1594
RESULT 15
BCK1 YEAST
ID BCK1 YEAST STANDARD; PRT; 1478 AA.
AC Q01369; P32894;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Serine/threonine protein kinase BCK1/SLK1/SSP31 (EC 2.7.-.-).
GN BCK1 OR SLK1 OR SSP31 OR LAS3 OR YJL095W OR J0306.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=92186847; PubMed=1545797;
RA Costigan C., Gehring S., Snyder M.;
RT "A synthetic lethal screen identifies SLK1, a novel protein kinase
RT homolog implicated in yeast cell morphogenesis and cell growth.";
RL Mol. Cell. Biol. 12:1162-1178(1992).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=92104496; PubMed=1840547;
RA Irie K., Araki H., Oshima Y.;
RT "A new protein kinase, SSP31, modulating the SMP3 gene-product
RT involved in plasmid maintenance in *Saccharomyces cerevisiae*.";
RL Gene 108:139-144(1991).
[3]
RN SEQUENCE FROM N.A.
RP MEDLINE=92107166; PubMed=1729597;
RA Lee K.S., Levin D.E.;
RT "Dominant mutations in a gene encoding a putative protein kinase
RT (BCK1) bypass the requirement for a *Saccharomyces cerevisiae* protein
RT kinase C homolog.";
RL Mol. Cell. Biol. 12:1172-1182(1992).
[4]
RN SEQUENCE FROM N.A.
RP MEDLINE=95174706; PubMed=7871887;
RA Miogsa T., Boles E., Schaaff-Gerstenschlaeger I., Schmitt S.,
RA Zimmermann F.K.;
RT "Sequence and function analysis of a 9.74 kb fragment of
RT *Saccharomyces cerevisiae* chromosome X including the BCK1 gene.";
RL Yeast 10:1481-1488(1994).
[5]
RN SEQUENCE OF 602-1104 FROM N.A.
RP Cusick M.E.;
RA Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
RL -!- FUNCTION: Serine/threonine protein kinase involved in a signal
CC transduction pathway that play a role in yeast cell morphogenesis
CC and cell growth. This pathway seems to starts by SMP3; then
CC involve the kinase PKG1 that may act on this kinase. BCK1 probably
CC phosphorylates MKK1 and MKK2 which themselves phosphorylate the
CC MEX1 kinase.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. MAP
CC kinase kinase kinase subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M84389; -; NOT_ANNOTATED_CDS.
CC EMBL; D10389; BAA01226.1; -.
CC EMBL; X60227; CAA42788.1; -.
CC EMBL; X77923; CAA54896.1; -.
CC EMBL; Z49370; CAA89389.1; -.
CC EMBL; Z49369; CAA89388.1; -.
CC EMBL; M88604; AAA21179.1; -.
CC PIR; S20117; S20117.
CC HSP; P24941; ICKP.
CC GeneOnline; 141709; -.
CC SGD; S0003631; BCK1.
CC GO; GO:0004709; F:MAP kinase kinase activity; IGI.
CC GO; GO:0000283; P:establishment of cell polarity (sensu *Sacch. . .*; IMP.
CC GO; GO:0007243; P:protein kinase cascade; IMP.
CC GO; GO:0007584; P:response to nutrients; IMP.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_thr_pkin_AS.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00069; Pkinase_i1.
CC PRINTS; PR00109; TIRKINASE.

DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TKc_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding; Phosphorylation.
FT DOMAIN 1175 1440 PROTEIN KINASE
FT NP_BIND 1181 1189 ATP (BY SIMILARITY).
FT BINDING 1204 1204 ATP (BY SIMILARITY).
FT ACT_SITE 1303 1303 BY SIMILARITY.
FT MOD_RES 1134 1134 PHOSPHORYLATION (BY PKC) (POTENTIAL).
FT VARIANT 1119 1119 T -> P (IN BCK1-11; ACTIVATION).
FT VARIANT 1120 1120 I -> K (IN BCK1-11; ACTIVATION).
FT VARIANT 1120 1120 I -> T (IN BCK1-16; ACTIVATION).
FT VARIANT 1146 1146 G -> V (IN BCK1-20; ACTIVATION).
FT VARIANT 1174 1174 A -> P (IN BCK1-20; ACTIVATION).
FT CONFLICT 59 59 F -> I (IN REF. 3).
FT CONFLICT 79 79 E -> V (IN REF. 2).
FT CONFLICT 264 264 A -> P (IN REF. 3).
FT CONFLICT 279 279 N -> I (IN REF. 3).
FT CONFLICT 703 714 RYEQTPSYVDYR -> STPKPRVITWTE (IN REF. 3).
FT CONFLICT 795 795 S -> A (IN REF. 3).
FT CONFLICT 802 802 L -> V (IN REF. 3).
FT CONFLICT 808 808 A -> S (IN REF. 3).
FT CONFLICT 903 903 T -> N (IN REF. 3).
FT CONFLICT 919 919 T -> N (IN REF. 3).
FT CONFLICT 960 962 ADA -> RDR (IN REF. 5).
FT CONFLICT 1086 1104 REVPPDSSVERFDGLNGK -> VPIAHTSSVYRMDLTVKIN
H (IN REF. 5).
SQ SEQUENCE 1478 AA; 164194 MW; D586C3A497A5B33 CRC64;
Query Match 6.8%; Score 299; DB 1; Length 1478;
Best Local Similarity 21.2%; Pred. No. 9.2e-08;
Matches 185; Conservative 131; Mismatches 338; Indels 218; Gaps 34;
QY 17 TDELSINKISATTD-----NSGTVQNI-----MMANNPEDWLSI-----LLKLEK 58
DB 678 TDVSVIVTSSSDITFTDFEHASGSGRRYPQTPSYVDYRVSNTNTPTBELNYWNIKEVLSHEE 737
QY 59 NSVPLSALLNKLIGRYQAEALPPDK-----YQONESFARIOV----- 98
DB 738 NA-----PKVFKTPKLELNLPDKGSKLNIPITENESKSFQVLRKDEGTEDF 789
QY 99 -----RFAELKAIQBPDDARDYFQMARANKKPAFVHISFAQFETISQGVNKKSK 147
DB 790 NHRRESPTYKPELAPKREAPK-PANTSPQRTLSLKQKPIRLVRASTKISRS--KRSK 846
QY 148 ---QLQKQAVERGA-----VPLEMLEIALRNLNLOKKQL----- 178
DB 847 PLPPQLLSPIEASSSPDLSLTSSVTPASTHVLIPOPYKGVANDVMRLKTDQDSTSTSPS 906
QY 179 LSEEEKONLSASTVLTAQESFGSGSLGHQNRNNSCDRSGQTTKARFLYGENMPPQDAEIG 238
DB 907 LKWKQKVNRNSTVSTNSIFYSPLLK-----RG----- 937
QY 239 YNSLRQNTKTKQSCFGRVFPVNLNSP---DCDVKTDDSVVPCFMKQKTSRSECRDLVV 295
DB 938 ---NSKRVSVSTSAADIFENDITFADAPPDFDSDSDSDS-----SSDDDIW 983
QY 296 PGSKPSGNDSCBLRNKLSQVNSHFKEPLVSDSEKSELSITITLKNKTESSLAKLEET 355
DB 984 SKKKTAPK-----TNNENKDEKSDNSSTH---SDEIFVDSQTD---KMERK 1025
QY 356 KEYQ-EPEVPSN-QKQWAKRKSECINQNSNHWQIPELARKVNTOKHTTFEPQV 413
DB 1026 MTFRESPEVYQNLKFFPRANLDKPIETEGIASPTSPKSLDLSLSPKQVASSRTEPSTP- 1084
QY 414 FSVSKQSPPISTSKWFD-----PKSICKTP-----SSNTLDDVMSCFRPVVKNDPPP 461
DB 1085 ---SRVPPDSSVEFIQDGLNGKPKPLNQAKTPKTKTIRIAHEASLAR---KNSVKL 1137
QY 462 ACQLSTPYQPACFQQQHQHILATPLQNLQVLASSANECISVKGRIVSYILKIGSGSS 521

Db	1138	KRQNTKQWGRMVEVTENHMSINKAKN-----SKGEYKEFAMWKG-----EWIGKSGFG	1187
Qy	522	KVFOVLN-----EKKQIYAIKYVNLLEADNOTLDSYRNEIAYLNKLOQHSDKIIRLYD	574
Db	1188	AVYLCNLVTTGEMMAVQVEVPKYSSQNEAILSTVEALRSEVSTLKDLDHLN--IVQYLG	1245
Qy	575	YEITDQYIYVME-CGNIDLNSMLKKKSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLK	633
Db	1246	FENKNTIYSLFLEYVAGSGSLIRMYGREDEPLIRHLTTQVLKGLAYLHSGILHRDWK	1305
Qy	634	PANFLI-VDGMLKLIDFGIANQMOPDTTSYVKDSQVGTVMYMPPEAIKDMSSSRENGKSK	692
Db	1306	ADNLLDQDGICKISDFGSRKSK-DIYSNSDMTMRGTVMFMAPEMV-----DTK	1354
Qy	693	SKISPKSDVWSLGCILYYMTYKTPFOQIINOISKLHAIIDPNHEIEFPDIPEKDL----	748
Db	1355	QGYSAKYVDIWSLGCIVLEMFAGKRPW----SNLEVAAMFKIGKSKSAPPIDETLPLIS	1410
Qy	749	---ODVLKCKLKRDPKORISIPELLAPHYVQI	777
Db	1411	QIGRFLDACFEINPEKPTANELLSHPFSEV	1442

Search completed: September 29, 2004, 16:16:46
Job time : 32 secs